

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2002, 23:38:05 ; Search time 10173.5 Seconds

(without alignments)  
5383.678 Million cell updates/sec

Title: US-09-171-553B-1

Perfect score: 3320

Sequence: 1 gaattcgcgagcgctgcgac.....aaaaaaaaaaaaaaaa 3320

Scoring table:

OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database:

Geneml:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pal:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hgo\_hum:\*  
31: em\_hgo\_inv:\*  
32: em\_hgo\_rod:\*  
33: em\_hg\_hum:\*  
34: em\_hg\_inv:\*  
35: em\_hg\_rod:\*  
36: em\_hg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3320	100.0	3320	6 A66551	A66551 Sequence 1
2	2682	80.8	8209	6 A66553	A66553 Sequence 3
3	2666	80.3	8196	6 A66552	A66552 Sequence 2
4	2321	69.9	8849	14 PEN133817	AJ133817 Porcine e
5	1769	53.3	8918	14 PEN133816	AJ133816 Porcine e
6	1680	50.6	7808	14 PER17013	AJ17013 Porcine e
7	1659	50.0	8763	14 PEN293657	AJ293657 Porcine e
8	1637	49.3	8750	14 PEN133818	AJ133818 Porcine e
9	1606	48.4	8918	14 PEN293656	AJ293656 Porcine e
10	1091	32.9	7333	6 AF038601	AF038601 Sus scrofa
11	1091	32.9	7333	6 AR130474	AR130474 Sequence
12	848	25.5	4402	6 AX052635	AX052635 Sequence
13	797	24.0	6076	6 AX052636	AX052636 Sequence
14	797	24.0	7362	6 AX052634	AX052634 Sequence
15	786	23.7	2462	6 AX002802	AX002802 Sequence
16	786	23.7	2462	14 PERENV1	PERENV1
17	746	22.5	6076	6 AX052647	AX052647 Sequence
18	704	21.2	8132	6 AR130475	AR130475 Sequence
19	570	17.2	4918	6 AX052637	AX052637 Sequence
20	557	16.8	8132	6 AF038600	AF038600 Sus scrofa
21	547	16.5	7873	6 AX052638	AX052638 Sequence
22	474	14.3	1980	6 AX052633	AX052633 Sequence
23	474	14.3	1980	14 AF130444	AF130444 Porcine e
24	438	13.2	2009	14 PEN288585	PEN288585
25	375	11.3	8060	6 AR130473	AR130473 Sequence
26	375	11.3	8132	4 AF038599	AF038599 Porcine e
27	321	9.7	3482	6 AX002804	AX002804
28	321	9.7	3482	14 PERENV2	PERENV2
29	285	8.6	716	14 AF163267	AF163267 Porcine e
30	208	6.3	927	14 SSU77599	SSU77599 Sus scrofa
31	170	5.1	942	14 MLVMT10	MLVMT10
32	166	5.0	166	14 AF138386	AF138386 Porcine e
33	166	5.0	166	14 AF138389	AF138389 Porcine e
34	161	4.8	810	14 AF033259	AF033259 Porcine e
35	158	4.8	522	14 AF033260	AF033260 Porcine e
36	158	4.8	927	4 SSU77600	SSU77600 Sus scrofa
37	119	3.6	615	14 AF163269	AF163269 Porcine e
38	97	2.9	1996	14 PEN288584	PEN288584
39	65	2.0	188989	10 AC074224	AC074224 Mus muscu
40	46	1.4	176506	2 AC079956	AC079956 Mus muscu
41	46	1.4	187393	2 AC084325	AC084325 Mus muscu
42	46	1.4	195423	2 AC087151	AC087151 Mus muscu
43	45	1.4	942	4 AF274705	AF274705 Sus scrofa
44	43	1.3	192465	2 AC091333	AC091333 Mus muscu
45	41	1.2	1591	10 AF161307	AF161307 Mus caroli

## ALIGNMENTS

RESULT 1  
LOCUS A66551  
DEFINITION Sequence 1 from Patent WO9740167.  
ACCESSION A66551  
VERSION A66551.1 GI:4538104  
KEYWORDS  
SOURCE  
ORGANISM

porcine endogenous retrovirus.  
porcine endogenous retrovirus  
Viruses: Retroviral viruses: Retroviridae; Mammalian type C  
retroviruses: 1-Mammalian type C virus group.

REFERENCE  
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.  
TITLE PORCINE RETROVIRUS  
JOURNAL Patent: WO 9740167-A 1 30-OCT-1997;  
Q ONE BIOTECH LTD (GB)  
FEATURES  
source  
1..3320  
Location/Qualifiers

/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

BASE COUNT 922 a 844 c 832 g 722 t  
ORIGIN

Query Match 100.0%; Score 3320; DB 6; Length 3320;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 gaattcgagcgcgctgcagatgctcttctgctgagatatacaacccactatgcca 60  
1 GAATTGCGGCGCGCTGCAGATGCTCTTCTGCTGAGATTAACCCCACTAGCCAA 60  
61 ccaatttgccttcgaatgagagatccagatgcggaagaacggcgacactctg 120  
61 CCAATTTGCTTGGAAATGAGAGATCCAGGTACGGGAACACGGGCGAGCTCCTGG 120  
121 acccgactgccccaaaggltcaagaactcccgaccatcttgaagaagccctaca 180  
121 ACCCGACTGCCCAAGGGTTCAAGAACTCCCGACATCTTTGACGAAGCCCTACACAGG 180  
181 gacctggccaacttcaagatccaacacctcaagtgaacctccctccagtaagtgtac 240  
181 GACCTGGCCAACCTCAAGATCCACACCTCAAGTGAACCTCTCTCCAGTACGTGATGAC 240  
241 ctgctcttgagcgagacccaacagagactgttaagaagtaagagacactactgctg 300  
241 CTGCTCTTGAGCGAGACCCAACAGAGACTGTTAAGAAGTAAGAGGCACTACTCTG 300  
301 gaattctgacctgagctacagagacctgttaagaagggccagatttgaagaagag 360  
301 GAATTGTCGACCTAGGCTACAGAGACCTGTGTAAGAAGGCCAGATTTGACGAGAGAG 360  
361 gtaacaacttgggtgtacagtttgcgggagcgagcgatgctgacggaagcggaag 420  
361 GTAACTGTAGTGGGTGTACAGTTTGGCGGGGCGGAGCATGAGTACGAGGCGAGAG 420  
421 aaaaactgtatccagataccgcccccaacaaagccaaagatgagagagattttgg 480  
421 AAAAAGTGTAGTGGGTGTACAGTTTGGCGGGGCGGAGCATGAGTACGAGGCGAGAG 480  
481 acagctggaatttgcagactgtgagatccggggttgcgaacctgaagcccaactac 540  
481 ACAGCTGGAAATTTGCAGACTGTGAGATCCGGGGTTTGGGACTTGAAGCCCACTTAC 540  
541 ccgctaaccaaaagaaaggggagatctcttgggtcctgagaccagaaagcattgat 600  
541 CCGCTAACCAAAAGAAAGGGGAGATCTCTGGGCTCTGAGACCAAAAGGCATTTGAT 600  
601 gctatcaaaaagccctgtagagcgacactgtctgacctccctgcgaactaataacc 660  
601 GCTATCAAAAAGCCCTGTAGAGCGACCTGCTGAGCCCTCCCTGACGTAACATAACCC 660  
661 tttaacctatgtatgagcgtlaaggagtagccgagagatttaacccaaccta 720  
661 TTTAACCTTATGTGATGAGCGCTAAGGAGTAGCCCGAGGATTTAACCCAAACCTTA 720  
721 ggaacatgagagagactgtgctacactgtaagaagaactgtaacctgaacagtgat 780  
721 GGAACATGAGAGAGACTGTGCTACTGCTCAAGAAACTTGAATCTTATACCAAGTGT 780  
781 tggccgtatgtctgaagcgtatcgacgtgtgacatgctggtcaagagcgtgacaa 840  
781 TGGCGCTATGTCTGAAGGCTATCGACGCTGTGACATGCTGATGGAAGCACTGACAAA 840  
841 ttgacttggagacagaataaactgtaagcccccaatgcatgtgagagaactgttgg 900  
841 TTGACTTTGGGACAGAAATATTAATAGCCCCCATGACANTGGAGAACTTGTGTGG 900  
901 cagcccccagacagatgagatgaccaaagcccgatgacccaactataaagcgtgtctc 960  
901 CAGCCCCAGACAGATGAGATGACCAAGCCCGATGACCCACTATCAAAAGCCTGCTTGC 960  
961 acagagaggttcaacttgcgtccacagcgctctcaaccctgacactcttgcctgaa 1020

1020 ACAGAGAGGTCACCTTGCCTCCACAGCGGCTCTCAAAACCTCCACTTCTGCTGAA 1020  
1021 gagactgtatgacagatgactcatgattgcataactatgtaaggaagctgagtc 1080  
1021 GAGACTGTATGACAGATGACTCATGATGATCCATCACTATTGATGTAGAGAGACTGGGTC 1080  
1081 cgaagagccttacaagataccgctgactgagagaagtgtcaacctggttactgaag 1140  
1081 CGAAGAGCCTTACAGATATACCGCTGACTGAGAAAGTGTCACTGCTTCACTGACGGA 1140  
1141 agcagctatgtgtggaaggttaagagatgtgtggcgagctgtgtgacggagccgc 1200  
1141 AGCAGCTATGTGTGGAAGGTAAAGAGTGTGGCGGAGTGTGACGGAGCCCGC 1200  
1201 acgatctggcccaagcgtccggaagaaacttcaacgcaaaaagcgtgactgagcc 1260  
1201 ACGATCTGGCCAGAGCGCTCCGGAAGAACTTCAACGCCAAAAAGCTGACTGATGGCC 1260  
1261 ctacagcaagcttgcggctgagcgaaggaatccataaactttaagagcagcag 1320  
1261 CTCACGCAAGCTTTCGGCTGCGCGGAAGAAATTCATTAACATTTATACGACAGAG 1320  
1321 tatgctttgagactgacagatcacagggccaacttaaaacaaaggggttgcctacc 1380  
1321 TATGCTTTGAGACTGACAGCTACCGGGCCATTTATAACAAAGGGGTTGCTTACC 1380  
1381 tcaagcagaggaagaaataaagaacaaaggaatctaaagctataagaagccttacc 1440  
1381 TCAGCAGGAGGAGAAATTAAGAAACAAAGGAAATTCATTAACCTTTATGAAGCCTTACAT 1440  
1441 ttgcaaaaagcttgcctatataactgtccctgagacacgaagaacaaagatctata 1500  
1441 TTGCAAAAAGCTTGCCTATATTAACACTGCTTGACACTGCAATGCAAAAGCAAGATCTATA 1500  
1501 tctagaaggaacagatgtgctgacgggttgcgaagcagcagcagcagctgttaacct 1560  
1501 TCTAGAAGGAACAGATGCTGACCGGGTTGCCAAGCAGCAGCCAGCGCTGTAACTT 1560  
1561 ctgctataatagaagacgccccaaagccccagaacccagaagacagatcaacctgaag 1620  
1561 CTGCTATATAGAAACGCCCAAGCCCCAGAACCCAGACCAACAGTACACCTTGAAGAC 1620  
1621 tggcagaagataaaaagaatagacagatctcttgaagactccgagggagactgtatacc 1680  
1621 TGGCAAGATATAAAAAGATAGACAGTCTCTGAGACTCCGAGGGAGACTGCTATACC 1680  
1681 tcaatggaaggaatctctgcccccaagaagaaggttaagatatgttccaaagatacat 1740  
1681 TCATATGGAGAGAAATCTCTGCCCAAAAGAAAGGTTAGAAATGTCCAAACAGATACAT 1740  
1741 cgtctaaccaacctagaactaacaacactgcagcagctgtgtcagaacaattcccttacc 1800  
1741 CGTCTAACCCACTTAGAATCAACACCTGCGAGAGTTGTGACAAACATVCCCTTATCAT 1800  
1801 gtctcagagctacagagatgagctgagctgtgtgtcaaacattgtgtcctgccaagt 1860  
1801 GTTCTAGAGGCTACAGAGATGCTGCTGACTGCGGTCAAAACATGTGTGCTGCCAGCTG 1860  
1861 gtttaatgtaatcttccagaataacctcagaagaagaactaaagggagacacccagc 1920  
1861 GTTAATGTCTAATCTTCCAGAAATACCTCCAGAAAGAGACTAAGGGAAGCACCCAGAGC 1920  
1921 gctcactggaagtgagacttcaatgaglaaagcggcgttaaaatagagaacaaatata 1980  
1921 GCTCAGTGGAAGTGAAGTCACTGAGTAAAGCCGGCTAAATATAGGAAACAAATATCTA 1980  
1981 ttggttttctgaacacatttcaagatgaggttagaggtctaccactaagaagaagact 2040  
1981 TTGTTTTTGTAGACACCTTTTCAGATGGGTAAGGCTTATCTTCAAGAAAGAGACT 2040  
2041 tcaacggtgtgctgaagaatactcggaggaatcttccaaagatttgaataactaaag 2100





RAMKOLREAVSGSDLOVPHRFQVDSVYVRHNRAGNLETRMKPVLVLTTPFAVKE  
 GIPWIIASHVKKPAPPDGSMRAEKTENPLKRLHRLVPISSNNNSPGO"  
 5620..7590  
 /codon.start=1  
 /product="VIRAL ENVELOPE PROTEIN"  
 /protein\_id="CA39349.1"  
 /db\_xref="GI:4538109"  
 /translation="MNPRTSRRLPLPRGEPKRLIPLSFASIAMFLTLTTTQASK  
 RLIDSNPHRLSLVLIIDPQGVNSTRGVARGTWPELHFCRLINPAKSTP  
 PNIVSYGFCCPEGEKEKTCGSGSPRRMSCTYSNDGMPKIISLDYKFSFV  
 SGPEKMKMLIKDKSCSPSDLDYLTISTEPKTKYSKVDELKNSFLVGGAG  
 STTLRIETGETEPVAMGPKVLEQGPALPBPNNIPVQSLSPDTQGPSNS  
 TTGLIPNTPNSBPVAVTKGQRLSLIGARQALNSDIPROATSCMLTSSGPYE  
 GMAKREKRVNTRKRNOCITWGSRNKLITLSEVSGTICISKAPRSHOICYSYVEOA  
 SENQIYKREPVSLTILAVMLIGTAVGVTSTVFNOSKDFVQVQIPRYIYREPVVLEDA  
 YRNNKREPVSLTILAVMLIGTAVGVTSTVFNOSKDFVQVQIPRYIYREPVVLEDA  
 LKESVNLBESITSLSEVYIQRNRLDLPLRGGGLCAALKECECPYDHSAITDSM  
 NLRKRLRRRRERERADQMEPEGNFNSPMWTLTSLALTPPLVLLLTIVGPECLINR  
 FVAVERVSAVQIVLVKQOYGLISGGEIDL"  
 FVAVERVSAVQIVLVKQOYGLISGGEIDL"  
 BASE COUNT 2168 a 2064 c 2154 g 1823 t  
 ORIGIN  
 Query Match 80.8%; Score 2682; DB 6; Length 8209;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2682; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 21 agatgctctctctgctgagatatacaccactagaccacttttgccttcgaatg 80  
 Db 2972 AGATGCTCTCTCTGCTGAGATTACACCCACTAGCACACTTTTGGCTTGCAATG 3031  
 QY 81 gagagatccaggtacggaagaacccgagcagctacacgtgacccagctgcccagaaggtt 140  
 Db 3032 GAGAGATCCAGGTACGGAAGAACCGGCGACCTCAGTCCGACCGACGCTGCCAAGGTT 3091  
 QY 141 caagaactcccgaccactcttgaagaagccctacacagagacactgcccacttcaagat 200  
 Db 3092 CAAGAATCCCGACCATCTTGGACGAGCCCTACACAGGACCTGGCCAACTTCAGAT 3151  
 QY 201 ccaacacccctcaggtacacccctcctcagtaagtgagacgttcttgcggaagac 260  
 Db 3152 CCAACACCCCTCAGGTACCCCTCTCCAGTACCTGGATACCTGCTTGGCGGAGGACAC 3211  
 QY 261 caaacaggaactgcttagaaggtacgaagcactactgcttgaatgtctgaactagctta 320  
 Db 3212 CAACAGGACTGCTTAGAAGGTACGAGGACTACTGCTGGAATGTCTGACCTAGGCTA 3271  
 QY 321 cagagcctctgctagaagaagccagatctgcaagagaaggttaacatacttggggtacag 380  
 Db 3272 CAGAGCCTCTGCTAAGAGGCGCCAGATTTCAGAGAGAGGTACATCTTGGGATACAG 3331  
 QY 381 tttagcgggagcggagcagatgctgacgagcagcaggaagaaactgtatgtccagatcc 440  
 Db 3332 TTTAGCGGGCGGCGAGCATGTGACGAGGACGAGCAAGAAACTGTGTGCTCAAGATAC 3391  
 QY 441 gggcccaacacagcaacaagtagaagattttggggaagcgtggattttgagact 500  
 Db 3392 GGGCCCAACACAGCAACAAGTAGAGATTGTTGGGACAGCTGATTTGACAGACT 3451  
 QY 501 gtgagtcgggggtttggagacttagcagcccaactctacacgcttaaccaagaagaag 560  
 Db 3452 GTGAGTCCGGGGTGGGACTTAGAGCCCACTACCTACCCGCTAACCAAAAAGG 3511  
 QY 561 gggatctctggtctctgagaccacgaagagcattgagatgcttcaaaaagccctgct 620  
 Db 3512 GGGATCTCTGGGCTCTGAGACACAGAGGCAATTTATCTTCAAAAAGGCGCTGCT 3571  
 QY 621 gaagcgaactgctgagccctccctgaagtaactaaccttacccttattgtgata 680  
 Db 3572 GAGCGAAGCTGCTGAGCCCTCCCTGAGCTAACCTTACCTTTATGAGATGA 3631  
 QY 681 gctaaagagtagccggaagattttaaaccacccctagaagacatggaagagactgt 740  
 Db 740 gctaaagagtagccggaagattttaaaccacccctagaagacatggaagagactgt 740

Db 3632 GCGTAAGGAGTAGCCCGAGAGCTTTTAACCCAAACCTAGACCATGAGAGACCTGT 3691  
 QY 741 tgcctactctcaagaagaactttagcttagccagtggttggccglatgctgaagcc 800  
 Db 3692 TGCTACTCTGCAAGAAGCTTGTATCTTACCGAGTGGTGGCCGATGTGGAAGGC 3751  
 QY 801 tatcgagcgtgtgccaactggttcaagaagcctgacaacttgaactttgggaagaaat 860  
 Db 3752 TATCGACACTGTGGCTACTGTGCAAGGACCTGTCAAAATGACTTTGGACGAAATAT 3811  
 QY 861 aactgtaataagcccccaatgcatctgtagaagaactcgttgcgagcccccagacagat 920  
 Db 3812 AACTGTAATAGCCCCCATGTGATGAGAACATCTGTCGGACAGCCCAAGACCGATGAT 3871  
 QY 921 gaacaaagcccgacatgaccacatacaaaagcgttcttcaagaagaggttcaacttcg 980  
 Db 3872 GACCAAGGCCCGCATGACCCTATCAAAAGCTGCTTCTACAGAGAGGTCACTTGGC 3931  
 QY 981 tccacaaagccgtctcaacccctgacactcttgcctgtagaagaactgtagaacagtgac 1040  
 Db 3932 TCCACCAAGCCGCTCTCAACCTGACCTCTTCTGCTGGAAGACATGATGAACCACTGAC 3991  
 QY 1041 tcatgattgcatcaactatgatttagaagaactggtgtccgaagaagccttacaagat 1100  
 Db 3992 TCATGATTGCTCATCACTATGATTGAGAGAGCTGGGTCGCAAGACCTTACAGACAT 4051  
 QY 1101 accgctgactgtagaagtgcttaactgctgctacatgacggaagcagctatgtgtgaaag 1160  
 Db 4052 ACCGCTGACTGAGAAATGCTTAACCTGCTTCACTGACGAGAGCTATGTGTGGAAGG 4111  
 QY 1161 taagaagatgctgtaggagcagtgatgtagcgggacccgacagatctgtgccaagact 1220  
 Db 4112 TAAAGAGATGCTGGGCGGAGTGTGACGGAGCCGACAGATCTGGGCGACACACT 4171  
 QY 1221 gcgggaagaaacttcaagcgaagaagcgtgagctatgctcgaagagcttggagct 1280  
 Db 4172 GCGGGAAGAACTTTCAGCGCAAAAGCTGAGCTCATAGGCGCTCACGACAGCTTTCGGCT 4231  
 QY 1281 ggcgaaggaagaaatccataaactttagcgaagcagagatgcttgccttgcagtcaca 1340  
 Db 4232 GGCGAAGGAAATCCATTAACATTTATACGACAGAGAGGTATGCTTGGACATGACACA 4291  
 QY 1341 cgtacacggggcactataaacaagaaggggttgccttactcagcgggaaggaataaa 1400  
 Db 4292 CGTACACGGGCACTTATTAACCAAGGGGCTTCTTACCTGACGAGGGAATAAA 4351  
 QY 1401 gaacaagaggaatctcagctttagaagccttcaatltgccaanaagcctagctat 1460  
 Db 4352 GAACCAAGAGGAATCTTAAGCTTATTAAGAACCTTACATTTGCAAAAAGCTTACTAT 4411  
 QY 1461 tatacactgtcctgagacatacaagaagccaagatctcactagaaggaagcagatgac 1520  
 Db 4412 TATACACTGCTGTGACATTCAGAAACCCAAAGATCTCATATCTAGAGGAAACAGATGAG 4471  
 QY 1521 tgaccgggttgcgaagcagcagccagcgtttaaactctgctcatalatagaagccg 1580  
 Db 4472 TGACCGGTTGCCAAGGAGGACCCAGGCTGTATTAACCTTCTGCTTATTAATGAAGAGCC 4531  
 QY 1581 caaagcccaagaaacacagacagcagctacacactagaagactggaagagataaagaag 1640  
 Db 4532 CAAAGCCCAAGAACCCAGACAGACAGTACACCTTAGAAGACTGCAAGAGATTAATAAATAT 4591  
 QY 1641 agacagctctctgagactccggaggggagactcttataactctataggaagaatcct 1700  
 Db 4592 AGACAGCTCTCTGAGACTCCGGAGGGAGCTGTATTAACCTCATATATGGAAGAAATCCT 4651  
 QY 1701 gcccacaagaagaggttagaatalgttcaacagatacactgcttcaacacactagaag 1760  
 Db 4652 GCCCCAAAGAAAGGTTAGATATGTCAACATATCATGTGTTAACCACACTAGAGAC 4711  
 QY 1761 taacactgtagcagattgttgaagaacatcccttatatattgttgcagctacagagat 1820  
 Db 4712 TAAACACCTGACACAGTGTGTGAGAAACATCCCTTATATGATGTGTGAGGCTACAGAGAT 4771



QY 861 aactgtaataagcccccatgcatctggagaaactgctgcgagcccccaagcatgat 920  
 DB 3798 AACTGTAATAAGCCCCCATGCACTGTGAGAAATCGTTGGGAGGCCCGCCAGACCGATGAT 3857  
 QY 921 gaaccaagcccgatgaaccatlatcaaaagcctgtcttcacagaagagtcacttcgc 980  
 DB 3858 GACCAACGCCCGATGACCCACATCAAAAGCCTGCTTCACAGAGAGGATCATCTTCCG 3917  
 QY 981 tccacagccgccttcaaccctgcacacttctgcctgaagaagactgtatgaacagtgac 1040  
 DB 3918 TCCACAGCCGCGCTCTCAACCCCTGCCTTCCTGCTGAAGAGACTGATGAACCGATGAC 3977  
 QY 1041 tcatgttgcctacactatgatgatgaagaactgtgggtcccgaaagaccttacagaat 1100  
 DB 3978 TCATGATGCGCATCAACTATTGATTGAGAGACTGGGGTCCCGAAGACCTTACAGACAT 4037  
 QY 1101 accgctgactgtgagaagctgaacactgttcaactgaagaaagactatgtgtgaaag 1160  
 DB 4038 ACCGCTGACTGAGAAAGTGTAACTGTTCCTGCTGACGAGAACACTATGTGTGAAGG 4097  
 QY 1161 taagaagatgctcgaggcgagctgtgtgagagggagcccgacagatctgtgcaagagct 1220  
 DB 4098 TAAGAGATGCTGGGGCGGCACTGTGTGAGAGCGGACCGACGATCTGGGCGAGAGCT 4157  
 QY 1221 gccggaaggaacttcaagcgcaaaagctgagctcatgtgcccacagcaagcttgcgct 1280  
 DB 4158 GCCGGAAGGAATCTTCAACGCGAAAGGCTGACATGCGCCCTCAGCAAGCTTTGCGGCT 4217  
 QY 1281 gccggaaggaatlcataaacaatlatacgaagcaagatgtgccttgcagtgcaca 1340  
 DB 4218 GCGCGAAGGAAATCCATAAACATTTATACGACAGCAGTATGCTTTGCGACGACACA 4277  
 QY 1341 cgtacaggggcatctatataaagaagggtgtgtctactctcagcagaaggaagaataaa 1400  
 DB 4278 CGTACAGGGGCACTATATAAAGAGGGGTTGCTTACTCAGCAGGAGGAGAAATATAA 4337  
 QY 1401 gaacaaaggaatcttaagactatagaagccttactatctgcacaaaaggctagctat 1460  
 DB 4338 GAACAAAGGAATTTTAAAGCCTATTAGAACGCTTACATTTGCCAAAAGGCTAGCTAT 4397  
 QY 1461 tatacactgtcttgagacatagaagaagcaaaagatctcatatctagaagggaaacagatg 1520  
 DB 4398 TATACACTGCTCTGACATAGAAAGCAAAAGCATCATATCAAGGAGGAGACCAAGTGAGC 4457  
 QY 1521 tgacgggttgcgaagcagcagccagcgtgtttaaactctgcacataatagaagcgc 1580  
 DB 4458 TGACCGGTTGCCAAGCAGGACGCCAGGCTTTAACTCTTGCCTATATATGAAGAACGCGC 4517  
 QY 1581 caaagccccaagaccagaagacgtacacccctagaagactgtgcaagagaataaaagat 1640  
 DB 4518 CAAAGCCCGACAGACCCAGACGACAGTACACCTTAGAAAGCTGGCAAGAGATATAAAAGAT 4577  
 QY 1641 agaccagttctctgagactcccgagggagactgtctactactatctatggaagaataatc 1700  
 DB 4578 AGACCAAGTTCTCTGAGACTCGGAGGGACCTGCTATACCTCATATATGGAAGAAATTCCT 4637  
 QY 1701 gccccaagaagaggtlagaatactgcacaagatatacatgcgtcttaaccacacccatagaac 1760  
 DB 4638 GCCCCACAAAGAAAGGTTAGATATGTGCCAAAGATATCATCGTCTAACCCACCTAGAAC 4697  
 QY 1761 taacaacctgacagttgttcagaaacatccctcatatctgtctcgtaggtctacagagt 1820  
 DB 4698 TAAACACCTTCAGCAGTTGTGTCAAGACATCCCTTATCATGTCTGAGGCTTACAGAGAT 4757  
 QY 1821 ggcgtgactcggtgtcaaaactgtgtgcacctgcacgtgtgttaagtataatcctccag 1880  
 DB 4758 GGCCTGACTCGGTGTCAAACTGTGTGCTTGCACCTGTGTTAATGCTATATCTTCCAG 4817  
 QY 1881 aatacctcagaagaagactaaagggaagccacacccagcgctcactgtggaagtgtgact 1940  
 DB 4818 AATACCTCCAGAAAGAACTAAGGAGGAGGACCAACGAGGCTACACTGGGAATGAGACTT 4877  
 QY 1941 cactgagtgtaagccgcttaaatcgggaacaataatctatctgtgttttctgtaaacctt 2000

DB 4878 CACTGAGTAAGCCGCGCTAAATACGGAACAAATATCTATTGGTTTGTGAGCACCTT 4937  
 QY 2001 ttcaagatgtgtgaagggtctatctacttaagaagaagactcaaccctgtgtgtgaagaa 2060  
 DB 4938 TTCAAGATGGTGAAGGCTTATCTACTTACGAAGAGACTTCAACCGGTGGCTTAAGAA 4997  
 QY 2061 aatacctgagaatatttccaagaatttgaatactaaaglaaaggttcagaacatgt 2120  
 DB 4998 AATACCTGAGGAATTTTCCAAAGATTGGAATACCTTAAGGTAAATAGGCTGACGAATGG 5057  
 QY 2121 tccagcttctgttgcacagtaagtcagagactgtgcgaagtatttgggattgtatga 2180  
 DB 5058 TCCAGCTTCTGTTGCCAGAGTAAGTCAGGAGCTGGCCAAAGATATGGGATTTGAGAA 5117  
 QY 2181 actgcatctgtacacagccccaagctcgaagaagctagaagaatgaataagaacat 2240  
 DB 5118 ACTGCACTTGTGCTACAGACCCCAAAAGCTCAGAGACAGTAGAGATGATAGACCAT 5177  
 QY 2241 taagaagaccttaccataatgacacagagactgtgcatlaatgatgtgacttcc 2300  
 DB 5178 TAAAGAGACCTTACCAAAATGACACAGAGACTGGCAATTAATGATGATGCTCTCT 5237  
 QY 2301 gcccttctgtctttaaaggtgaagaacacccctgtgaagatttggctgtaccctatga 2360  
 DB 5238 GCCCTTGTGCTTTTAAAGGTGAAGAACACCCCTGACAGATTTGGCTGACCCCTATGA 5297  
 QY 2361 attgctctacggggagcccccccgctgtgcagaatattgctcttgacataagtgctatgt 2420  
 DB 5298 ATTGCTCTACGGGGAGACCCCGCTGGACAAATTTGCTTGCACATPATGCTGATGT 5357  
 QY 2421 gctgcttcccaagccttgttctcctagctcagaagcgtcaggtgtgtgagcgagc 2480  
 DB 5358 GCTGCTTCCACAGCTTTTGTCTGTAGCTCAAGGCGCTCGATGGGTGAGGACAGCAGC 5417  
 QY 2481 gtggaagcaactccgggagggcctactcaggaagaagactgtcaagttccaactgttcca 2540  
 DB 5418 GTGGAAGCACTCCGGAGGCGCTACTCAGAGAGAGACTTGCAAGTTCCACATCGCTTCCA 5477  
 QY 2541 agttgagattcagctatgtlagaagcagcgtgtgagaagaacccctcgagactcgttgaa 2600  
 DB 5478 AGTTGAGATTTCAAGTATGTTAGAGCGCCACCGTCGAGAAACCTGAGACCTGAGAA 5537  
 QY 2601 ggaacctatctgactcttgcacacaccaaaggtgtgaaagtcgaagaatccac 2660  
 DB 5538 GGGACCTTATCTGTACTTTTGAACACACCGCTGTGAAGTCGAAGGATCCCAAC 5597  
 QY 2661 ctgagatcatgcatccacgttaagc 2686  
 DB 5598 CTGATTCATGATCATCCACGTTAAGC 5623

RESULT 4  
 PEN133817 8849 bp DNA VRL 07-JAN-2000  
 LOCUS Porcine endogenous retrovirus type C proviral gag, pol and env  
 DEFINITION genes and LTR (claus A, clone 42).  
 ACCESSION AJ133817  
 VERSION AJ133817.1 GI:6688947  
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus  
 Viruses; Retroid viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 8849)  
 Cauderna, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and  
 Toenjes, R.R.  
 TITLE Molecular characterization of Human-tropic and  
 Replication-competent Porcine Endogenous Retroviruses  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 8849)  
 AUTHORS Toenjes, R.R.



Db	4279	TATCGAGCTGTGGCATACTGTGTCAGAGACGCTGACAAATTTGACTTTGGACAGANAT	4338
Qy	861	aactgtaataagcccccatgcatcttgtagaacaatcgttcgcaagccccagacagatgta	920
Db	4339	AACCTGTAATAGCCCCCATGTCATTTGGAGAAACATCTTGGCAGACCCCCAGACGATGAT	4398
Qy	921	gaccacagcccgatgacacacatcaaaagcctgtcttccaaagagaaggtcaattgcg	980
Db	4399	GACCAACGCCCGCATGACCCATATCAAAAGCCCTGCTTCACAGAGAGGGTCACTTTCCG	4458
Qy	981	tcacacagcccgcttcaaacctgtccacttctgccttgaaagagacgatgaaacagtgac	1040
Db	4459	TCACACAGCCGCTCTCAACCTTCCACTCTTCTGCTGGAAGACATGATGAACCAATGAC	4518
Qy	1041	tcattgattgcacacacatctgattagaggaactgggtccgcaagaccttaacaacat	1100
Db	4519	TCATGATGTGCATCACTATTTGATGAGAGACTGGGGTCCGCAAGACCTTACACACAT	4578
Qy	1101	accgctgactggaagagtgctaaacctgttctactgaaggaagcagctatgtgtgaag	1160
Db	4579	ACCGTACTGAGAAAGTGTACCTGCTACCTGACGCAAGCAGCTATGTGTGGAGAG	4638
Qy	1161	taagagatgctggggcgagctgtgtgagcgagaccgacagatctggccagagcct	1220
Db	4639	TAAAGAGATGGTGGGGGCGAGTGGTGGAGCGGACCCGACATCTGGCCAGACCT	4698
Qy	1221	gccggaaggaacttcagcgcaaaaggctgagctcagctcagccctcagcaagcttgcgct	1280
Db	4699	GCCGGAAGGAATTCAGCCCAAAAGGCTGAGCTCATGGCCCTCAGCCAAAGCTTTGGGCT	4758
Qy	1281	ggccggaaggaatccataacattataagagcagagatgagcttggcagctgcaca	1340
Db	4759	GGCCGAAGGGAATTCATAAATTTATACGACACAGGATGAGCTTTGGCAGCTCACA	4818
Qy	1341	cgtaacaggggcatctataacaaagggtgtgcttaccctcagcgagggagaaataa	1400
Db	4819	CGTACAGGGGCCATCTATTAACAAGGGGCTTGTACTCTCAGAGGAGGGAATTA	4878
Qy	1401	gaacaaaggaatcttaagcctattagaaccccttaacattgcccataaagagctagcat	1460
Db	4879	GAACAAAGAGGAATTTCTAAGCTTATAGAGCTTACATTTGCAAAAGGCTTACTAT	4938
Qy	1461	tataactgtctggaactaagaagccaagaatctcatactagaagggaacagatgac	1520
Db	4939	TATACACTGTCTGGAATAGAAAGCAAGATCTCATATCTAGAGGGAACAGATGGC	4998
Qy	1521	tgacgggttgcaagcagcagccagcgtgttaacctctgtcctataatagaacagcc	1580
Db	4999	TGACGGGTTGCCAAGCAGGACGCCAGGCTGTAACTCTGCTATATATAGAAAGCC	5058
Qy	1581	caaaagcccaagaccagacagcaagtaacccctagaagactggcaagagataaaagat	1640
Db	5059	CAAAAGCCCAAGACCCAGAGACAGTACACCTAGAAAGACTGGCAAGAGTAAAGAAAT	5118
Qy	1641	agaccagttcttgagactccgaggggagctgcttaactcaatctatggaagaaatcct	1700
Db	5119	AGACCACTTCTTGAGACTCCGGAGGGAGCTCTACTACTCATATGAGGAAGAAATCCT	5178
Qy	1701	gccccacaagaagggttagaataatgtccaacagatacatcgtctaaacccactaggaac	1760
Db	5179	GCCCCACAAGAAGGGTTAAATATGTCCAACAGATACATGCTGTACCCACTAGGAAC	5238
Qy	1761	taaacactcagcagagtgtgtgaagaatacccttactcatgttcttgaagttacagaagt	1820
Db	5239	TAAACACTCTGACAGTGTGGTCAAGACATCCCTTATCATAGTTCTGTGGGTACAGAGAT	5298
Qy	1821	ggtcactcgtgtgtcaaaatgtgtgcccctccagctgttcaatgcttaactctccag	1880
Db	5299	GGCTGACTCGGTGTCAAACTATGTGTGCCCTGCGCAAGCTGTTAATGCTAATCCTTCCAG	5358
Qy	1881	aatactccaggaagagactaaagggaagccacccagcgctcactggaagtgagctt	1940
Db	5359	AATACCTCAAGGAAGAGACTAAGGGGAAGCCACCCAGGCGCTCAGCTGAGGAAGTGGACT	5418
Qy	1941	cactgaggtlaaagcccgcttaataacgaaacaaatattatgttctttagaacctt	2000
Db	5419	CACGTAGGTAAAGCCCGCTAAATATAGGAACAAATATCTATGCTTTTGTAGACACCTT	5478
Qy	2001	ttcaagatgggttagaggttcttcactaaagaagagactcaacgcttgggtgagtaa	2060
Db	5479	TTCAAGATGGGTAGAGGCTTATCTTACTTAAGAAAGAGACTTCAACCGTGGGTGA	5538
Qy	2061	aatactgaggaatcttccaaagattggaatacctaaggtaaaggtlcaagaacatg	2120
Db	5539	AATACTGGAGGAATTTTCCGAGATTTGGAATACCTAAGGTAAATCGGGTCAGCAATGG	5598
Qy	2121	lccagcttgcgttgcacaglaagtcagggactggccaagatattgggattgtgaa	2180
Db	5599	TCACACTTTCGTTGCCAGGTAAAGTCAAGGAGCTGCCCCAAGATATTTGGGATTTGGA	5658
Qy	2181	actgcatgtgatacacaacccaaagctcagcaggttaagaggaatgaaatagaacat	2240
Db	5659	ACTGCATGTGATACAGACCCCAAGCTCAGAGAGGTAGAGAGATGAATTAACCAT	5718
Qy	2241	taaaagaccccttaacaaattgacacagagactggcattatgtattgagctcct	2300
Db	5719	TAAAGAACCCCTTACCAATTTGACCAAGAGACTGGCATTAATGATGTGATGCTTCCCT	5778
Qy	2301	gcccttctgcttctttagagtgaggaacaccccttggaagcttgggtgacccctatga	2360
Db	5779	GCCCTTTGCTTTTAAAGGTAGAGAACCCCTGACAGTTTGGGCTGACCCCTATGA	5838
Qy	2361	attgcttcaagggagaccccccgttgcagaatctgcttgcacatagtgctatgt	2420
Db	5839	ATTGCTTAAGGGAGGCCCCCCCGCTGTGTGAATTTCTTCTGTACATATGCTGATGAT	5898
Qy	2421	gctgtcttccagccttgcctccttagcctcaagggcgtcagtgagtgagcgagc	2480
Db	5899	GCTGCTTCCAGCCTTGTCTTCTTGAAGGCTCAGGCGCTGAGTGGGAGGCAAGCAGC	5958
Qy	2481	gtggaagaagctccggagggccttactaggg---aggaacttgcagatctcaatcgctt	2537
Db	5959	GTGGAAGCAGCTCCGGAGGCGCTACTCAGAGAAAGAACTTGCAAGTTCCACATCGCTT	6018
Qy	2538	ccaagtctgagatctcagctatgttagacgcacacgcttgcaggaacactcgaactcggtg	2597
Db	6019	CCAACTGGAGATTCAGTGTATGTATGAGACGACCGTGCAGGAACCTCGAGACTGGTG	6078
Qy	2598	gaagggaactatctcgtactcttgcacacacacacagcgtctgaaagtcgaagaatccc	2657
Db	6079	GAAAGGCCCTTATCTGCTTGTGACCAACACAGGCGCTGGAAGTCAAGAAATCCC	6138
Qy	2658	cacctgagatccatccacagcttaagccgagcgccacccctccgattcgggggtgaaagc	2717
Db	6139	CACCTGGATCCATGCAATCCACGTTAAGCCGGCGCCACTCCGATTTGGGGTGAAGGC	6198
Qy	2718	cgaaaagactgaaaaatcccttaagcttcgcctccatcggtgtgttcttactctgcaa	2777
Db	6199	CGAAAAAGACTGAATAATCCCTTAAGCTTGCCTCATGCGCGGTGCTTACTGTGCA	6258
Qy	2778	taacctcaagttaatgtgaaagccttgggaagccggaaccccaataaaccttat	2837
Db	6259	TAACTCTCAATTAATGATTAACGCTTGTGCAAGCCGCAACTCCATTAACCTTAT	6318
Qy	2838	ctctcaactggttactactgactcgggtacaggtatataataataacagcactcaagggt	2897
Db	6319	CTCTCACTCGTACTTACTGACTCGGTAGCAGTATTAATTAACAGCACTCAAGGG	6378
Qy	2898	aggtcccttgggaactgtgtgctgaaatataatgtcctgcctcgatcaglaatccctg	2957
Db	6379	AGGCTCCCTTGGGGAGCTGTGGCTGAATTAATGTGTGCTTGCATCAGTAATCCCTG	6438
Qy	2958	gtctcaatgacagggcaaccccccgatgtactcgtgtgaggttaaggttttagcttgc	3017
Db	6439	GTCTCAATACCAAGGCCACACCCCGGATGTACTCGTGTCTTACGGGTTTACGTTTGCC	6498

QY 3018 caggaccaccaataatgaagaatlatgtygaatcctcaggaattctcttgcaagcaat 3077  
 |||||||  
 Db 6499 CAGAGCCCCCAATATGATGGAATATTGTGGAATCTCAGAGATTTCTTTTCAAGCAAT 6558  
 |||||||  
 QY 3078 gggagctgataacttaataatgaggaatggaatgagcaagctctcagcaagaagag 3137  
 |||||||  
 Db 6559 GGAGCTGCAATACCTTAAATGATGGAATGGAATGCGACGCTCTCTGCAAGACAGAG 6618  
 |||||||  
 QY 3138 taagtaactcttctgtaaacatccctacagttataatcaattatgagcaatgga 3197  
 |||||||  
 Db 6619 TAAGTACTCTTTTGTGTAACATCCCTACCAAGTTATATCAATTTATATGCGCATGGGA 6678  
 |||||||  
 QY 3198 gatggaagaattggaacagcggtacaaagaatgataagcaataaagctctc 3257  
 |||||||  
 Db 6679 GATGGAAGAATGTCGAACAGCGGCTACAAAGATGTACGAATTAAGCAATTAAGCTGTC 6738  
 |||||||  
 QY 3258 attcgtagaactgaattacttaaaataagttcaat 3295  
 |||||||  
 Db 6739 ATTGCTTAGACCTGATTAATTAAGTTCACT 6776  
 |||||||

RESULT 5  
 LOCUS PEN133816 8918 bp DNA VRL 07-DEC-2000  
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env  
 genes and LTR (class B, clone 33).  
 ACCESSION AJ133816  
 VERSION AJ133816.1 GI:6688945  
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol  
 protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM Porcine endogenous retrovirus.  
 Viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 REFERENCE 1 (bases 1 to 8918)  
 Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.  
 Establishment and characterization of molecular clones of porcine  
 endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)  
 JOURNAL MEDLINE  
 REFERENCE 2 (bases 1 to 8918)  
 Tonjes, R.R.  
 Direct Submission  
 Submitted (04-MAY-1999) Tonjes R.R., Medical Biotechnology,  
 Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225  
 Langen, GERMANY  
 FEATURES  
 Location/Qualifiers  
 source 1..8918  
 /organism="porcine endogenous retrovirus"  
 /strain="type C"  
 /db\_xref="taxon:61673"  
 /clone="33"  
 /note="grown in PK15 (ATCC NO. CCL-33) cell line"  
 1..707  
 LTR misc\_feature 545  
 /note="cap site"  
 primer\_bind 710..727  
 /note="gly4"  
 gene 1154..2725  
 /gene="gag"  
 1154..2725  
 /gene="gag"  
 /function="retroviral capsid"  
 /product="gag protein"  
 join(1154..2725,2729..6313,6189..8162)  
 /note="viral polypeptide"  
 /codon\_start=1  
 /protein\_id="CAB5339.1"  
 /db\_xref="GI:6688946"  
 /db\_xref="SPTREMBL:Q9Q1X5"  
 /translation="MGQVTTPLSLTLDHMTVEVSRANHLNVQYKKGPMQTFCASEMP  
 NFDVGPSEGTFNSRIILAVKAILFQTPSSHPQDEYIILTWQDLADPPWPKWKLIN

KPKKPEPRILALGKKNKSAEKEVPSRIYELIEPPTWPEPPEPPYPAQAGVRC  
 PSAPGAPVVEDEPAAGTNRSGATPERDELAILEPRTYGPMPGOLQLOYPFSS  
 ADLYNMKTNHPFSEDEPQRLTGLVESLMSHOPMDCCOQLOLPTFTREERILLER  
 RKNVPCADGRPTOLONEIDMFPLTRPMDMDNATGRESLITIRALVATGIRGSRPR  
 TNLAVREVMQGNPPPSVTELERMEARRRPTPDPTSEAKASVALAIGSALDIR  
 KLIQRIEGLQELDLIVREKERYRRETEEEKQKREEREEREDRROCKNLT  
 KILAAVGEKSSREREDEFRIRSGPRSGNIMNNTPLDKQCAVCEKGMANCK  
 KNGKPKVTLLEEDKGRSGSDPLPEPVTTLKVGQPEVTLVDGASAKKAOICREPVYICIS  
 KEKSWMGATGGOROYPWTRRTVDLAVGRTSHSELVIPCPEVTLDRDLITLKGAOI  
 SPEQGRPEVSNMKNPITVLTLOLDEVERLSPQKPPQDIOISMLQEPQMAETKAGG  
 LAKQVPROVITOLKASATPVSVKQPLSREAREGIMPVORLQOGLILVQSPMNPPL  
 LPKRVGINDRIPVQDLKVNKRVQDHPVNPVNLISLAPERNMVTYLDLKAFF  
 CLRHPTSPLEFAFEMRDPQGTGOLTWRLPQGFNSPTIDEALIRPLANERIOH  
 POVTLQYVDDLALAGATKQDLGKALILELSDLGFRASAKKAOICREPVYICIS  
 LSGQRMLEAKRKTYVOIPAPVTAKQVRELTGAGCGRMTGEPATLAPVPLKRE  
 KGEFMAPEHOKAPDAIKKALLSPALADPVKPTPTLYDEKGVAKGVLVLTLPW  
 RPPVAVYLSKIDLPVASGMPVCLKILAAVAILVADADLITLQONITVYAPHALEIVRO  
 PDRMNTAKRNTYQSLTLREYTFAPPALNPAVTLLEETDEPVRHDCQLILEETG  
 VKKDLTDLPLTGEVLTWETDSSYVEGKRMAGAAYVDGTHIIMASLPGTSQAKAE  
 LMALTOALRLAEKKSINITYDSRYAFATAVHGAITYKQRLITLSAGREIKREKETSIL  
 LEALHLPKRLALIHCPGHOKAKDLISRGNDMPDVAQOANVNLPIETTPAPEPR  
 ROYTLEDMQELKKTIDQSEPECTCTYSKEILPHKEGLETVQOHLRLTGLTKHLQ  
 OLVRTSPETHVRLPQVADSVKHCVPQOLVANANPSRLIPKRLRSGHPGAMVDETE  
 VKPAPYGNKYLIVLVEFSGWVEAYPRKKESTVVAKKILEETPFEGIKVIGSDNG  
 PAPAQVSQGLAKTIGIDMKLHCAYRQSSGOVERMNRITKETLTKLTETGINDMIA  
 LLPVLEFRVNTPCQFGLTPELILYGGPPPLVEATVSHSADVLISQPLSRKALBMY  
 KORAMKOLREAYSGEGDLOVPHROVDVSVYRHRNAGNETWKKPVTLYLTTPAY  
 KVEGISTWIIHSHYKLAIPDPSGGRAKTENPILRLHRLVPIYNNNSPQHLPTLW  
 RHLPTRGGERRLKPIPLSFASIAFELTTLTPASSKRLLIDSSNPRPLSLWLIIDP  
 DTGTVANSTRGVAPRTWPELHFCLELIPAVKSTPNPNVRSYGYCCPGTEKERYC  
 GDSESEFCRMSCVTSNDQWMPKPTISLQDVKFSFVNSGKRYKVKLKKDKCSFSD  
 LDYLKISTFEKQENIOWINGSMQIVYKKGSGSLTILRLTIPGTETPVAVG  
 PDKVLAEQPEALPEPNIPLPOLTSIRPTIOTNSNGTGLIPTPTPNRSQVPKAT  
 GQRLFSLIOGAFQAINSTDPTDSCMLCISSPPTBYEMAKGKNVIREHNDQTM  
 GSRKILTLTVESSGCTICGAPPSHOLCTSTVYEQDASNDQNLVGYRMAACNGL  
 TPCVSTVFNOSKFCVWQIVPRVYHPEPVYIDEVDEYNNPKREPVSLTLAVLG  
 LGTVAGVGTGIALITGPQOLEKGLGLHAAEMEDIALAESVNLSESLTSEVYL  
 ONRGVLDLPLEGELCALKECCFVHSDGAIIRMSKALRRLERLRERREEDQW  
 FEGWENSPMWTTLALSALGPLVLLLLLVGSCILINRPAVYREVSVAQIMVLRQO  
 YQGLISQGETDL"  
 2729..6313  
 /gene="pol"  
 /function="retroviral protease and polymerase"  
 /product="Pol protein"  
 2729..6313  
 /gene="pol"  
 6189..8162  
 /gene="env"  
 /function="retroviral envelope"  
 /product="Env protein"  
 6189..8162  
 /gene="env"  
 8212..8918  
 /note="env"  
 LTR 8212..8918  
 polyA\_signal 8811..8816  
 BASE COUNT 2327 a 2245 c 2334 g 2012 t  
 ORIGIN

Query Match 53.3%; Score 1769; DB 14; Length 8918;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 2219; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 64 cttttgctgaaatgagagatccaagtacggaggaagaacggagcactggacc 123  
 |||||||  
 Db 3581 CTTTITTCCTTCGAATGGAAGATCCAGGTACGGGAAGACGGGACGCTGACAC 3640  
 |||||||  
 QY 124 cgactgccccaaaggttcaagaactcccgacacatttgagcaagagcctacaaggagc 183  
 |||||||  
 Db 3641 CGAGTGGCCCCAAGAGGTTCAGAAAGCTCCCGACCATCTTTCAGAGGAGCCCTACACGGGAGC 3700  
 |||||||  
 QY 184 ctggccaacttcaagatccaacaacccctcagtgacccctcctcagtaagtgaatacctg 243  
 |||||||  
 Db 3701 CTGGCCAACTTCAGATTCACACACCTCAGTGACCCCTCTCACTACGTGATGACCTG 3760  
 |||||||









QY 801 tatcagactgtgcccatactgtgtaaggaagcttacaattgacttggagacat 860  
 |||||||  
 Db 3424 TATCGAGCTGTGGCTATCTGGTCAAGAGCGCTACAAATTTGACCTTGGAGCAGAATAT 3483  
 QY 861 aactgtataagcccccatgcatgtgagaacacgttcggaagcccccaagccatgcat 920  
 |||||||  
 Db 3484 AACTGTATATAGCCCCCATGATGTGAGAACATCTGTGGCAGCCCCCAGACCATGAT 3543  
 QY 921 gacaaagccgcaagagaccactatcaaaagctgttctcaagagaaggttcaattcgc 980  
 |||||||  
 Db 3544 GACCAACGCCCGCATGACCCACTATCAAGCCCTGCTTCTCAAGAGAGGCTTCTTCCG 3603  
 QY 981 tccacagccgctctcaaccctgcacactcttctgcttcaagaagatgtaacacagtgac 1040  
 |||||||  
 Db 3604 TCCACACGCGCTCTTACACCTGCGCACTTCTGCTGAAGAGATGATGAACCAAGTAC 3663  
 QY 1041 tcatgttgcatcaactatgtatgtgagaagactgtgggtccgaaagaccttaacagat 1100  
 |||||||  
 Db 3664 TCATGATTTGCCATCAACTTATGATGAGAGAGACTGGGGTCCGCAAGGACCTTACAGACAT 3723  
 QY 1101 accgctgactggaaggtgttaacctgttcaactgttcaaggaagacatgtgtgtggaag 1160  
 |||||||  
 Db 3724 ACCGCTGACTGGAGAGTGTCTAACCCTGTTCTACTGAGGAGACACTATGTTGTGGAAG 3783  
 QY 1161 taagaggaatgctgagggcgagtgagtgagcggaacccgcaagatcttgagccagagct 1220  
 |||||||  
 Db 3784 TAAGAGGATGGCTGGGGCGGGTGTGTGAGCGGAGCCACAGATCTGGCGCAGCAGCT 3843  
 QY 1221 gccggaagaaacttcaagcgcaaaagctgagctcatgagccctcaagcaagctttgcgct 1280  
 |||||||  
 Db 3844 GCCGGAAGGAATCTTCAAGCCCAAAAGGCTGATGATGAGCCCTCAAGCAAGCTTTGCGCT 3903  
 QY 1281 gccggaagaaatccataacaattatagcagcaagatgtgcttctgcaagcgca 1340  
 |||||||  
 Db 3904 GCGCGAAGGAATCCATTAACCTTTATAGGAGACAGACAGATGTGCTTGGCAGCAGCA 3963  
 QY 1341 cgtacaggggcatctataaacaaggggtgtgttacttccagcaagggaggaataa 1400  
 |||||||  
 Db 3964 CGTACAGGGGCTATTAATAAAGGGGTTGCTTACTTACAGAGGAGGGAATATA 4023  
 QY 1401 gaacaaaggaatcttaagccttatagaagccttaactgtgcaaaagctagcat 1460  
 |||||||  
 Db 4024 GAACAAAGAGGAATCTTAAGCCTATTAGAGCCTTACATTTGCCAAAGAGCCTTAGCTAT 4083  
 QY 1461 tatatcgtccttgcacatcaagaagccaaagatcctatatactagaaggaacagatgac 1520  
 |||||||  
 Db 4084 TATACACTGTCTGAGACATAGAAAGCCAAAGATCATATCTAGAGGAGACAGATGAC 4143  
 QY 1521 tgacgggtgtgcaagagcagccagagctgttaaccttgcctataatagaagcc 1580  
 |||||||  
 Db 4144 TGACCGGGTTGCCAAGAGGACGCCAGCTGTTAACCTTCTGCTTATATATGAAACGCC 4203  
 QY 1581 caaagcccaagaaacccagagcagatcaacccctagaagaactggcagaagataaagaat 1640  
 |||||||  
 Db 4204 CAGAGCCCCAGAACCCAGACGACATGACACCTAGAGAGCTGGCAAGAGATATAAAGAT 4263  
 QY 1641 agacagcttcttgagacttcggaagggagactgctatacctatataatgggaagaaatcct 1700  
 |||||||  
 Db 4264 AGACAGTCTCTGAGACTCCGAGAGGAGACTCTTATACCTCATATGAGGAAGAAATCCT 4323  
 QY 1701 gccccaagaagaggttataatatagtccaagaagataactgttcaaccactagaagac 1760  
 |||||||  
 Db 4324 GCCCCCAAGAAAGAGGTTAATAATATGTCCAACGATATCTGCTTCCACCTCTGGAAC 4383  
 QY 1761 taacacactcagcaagttgtcagaacatcccttctcatgttctgaagctcaggaagat 1820  
 |||||||  
 Db 4384 TAAACACTCTGACAGATTGTTGTCAGAACATCCCTTATCATGTTCTGAGGCTTCCAGGACT 4443  
 QY 1821 ggtcactcgtgtgtaaaatattgttgcctgtccagctgtgttaatgtaactcctcag 1880  
 |||||||  
 Db 4444 GGCTGACTCGTGTGTAACATATGTGTGCCCTGCCAGCTGTGTAATGTAATCTTCCAG 4503  
 QY 1881 aatcctcccaagaaagagactaagggaagccacccagcgctcactgtggaagtgtgactt 1940

|||||||  
 Db 4504 AATACCTCCAGAAAGAGACTAAGGGAAGCCACCAAGCGCTCAGTGGGAATGCACTT 4563  
 QY 1941 cactgaggtaaagccgcttaataaggaacaaatattatgttttttgaagcactt 2000  
 |||||||  
 Db 4564 CACTGAGTGAAGCCGCTTAATFACGAACAAATATCTATGTTGTTTGTGAGACACTT 4623  
 QY 2001 ttcagatgtgtagagcttactcactaagaagaagactcaacgctggtgtgctaa 2057  
 |||||||  
 Db 4624 TTCAGATGGGTAGAGGCTTATCTTACAGAAAGAGACTTCAACCGGTGGCTTAA 4680

RESULT 7  
 PEN293657  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 REFERENCE  
 2 (bases 1 to 8763)  
 AUTHORS  
 Toenjes, R.R.  
 Direct Submission  
 Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology,  
 Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,  
 GERMANY  
 FEATURES  
 source  
 location/Qualifiers  
 1..8763  
 /organism="porcine endogenous type C retrovirus"  
 /proviral  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:70540"  
 /clone="213"  
 /cell\_line="PK15"  
 /note="Class B"  
 1..630  
 467  
 /note="Cap site"  
 633..650  
 1077..2651  
 /gene="gag"  
 1077..2651  
 /gene="gag"  
 /codon\_start=1  
 /product="group specific antigen"  
 /protein\_id="CAC38625.1"  
 /translation="MGQVTTPPLSLDHWTEVRSRAHNLSVOYKKGPWOTFCASEMP  
 PEDVGPSEGFENSEIILAVKAIIFGPGSPDQEPYIILWDLDEPPVWPKPMLN  
 KRKRGPRILALGKKNHSAKVPSPGTYEIEIEPTWPPROPVPPYPAQAGAVN  
 PSAPPGAYVEBPAPAGTSSRGATPERPDELAIILPTGPPMGGLOLPWPPSS  
 ADLVWKNTHPPESDPORLTLGLVESLMFSHPPTWDDCOOLLTLFTTEREKILIEA  
 RNKVPADGRPTOLNEIDMGFPILTRPGWDNTAEGRESLKIYKQALVALRGSRRP  
 TMLAVREVMQPNPSPVFLERLMEAFRRPTPDPSKQKASVALATIGQSLDIR  
 KKLRLREGLOEALRDLYRAEKVYVYRRETEEEKEREREREERERDRROEKNL  
 KTLAAVVEGSSRSRERDRFRKIRSPRQSGNLGRPTLIDKQCAVCKEKHMANCPK  
 KKNKPKVLALEEDD"  
 2652..6239  
 /gene="pol"  
 <2652..6239  
 /gene="pol"  
 /note="complete pol ORF, gag-pol is expressed as a

```
polyprotein"
/citation="[1]
/codon_start=1
/product="polymyrase"
/protein_id="CAC39626.1"
/db_xref="GI:14275843"
/translation="GRKSDPLPEPRVTLKVEGQVPEFVLDGAEHSLLOPLGKLE
KSMWAGATGQROYPTWTRRTVLAAGVTHSVFVTEPCPVLLGRLLLRKMAQVSE
EGREVSNNKPIVLTLOLDEERLXSPQVKDDOISMLRPFQANAEATGKMLA
KQVPOVIOKASATPVSVROYPLSREAREGIMWYORLLODGLIVVQSPMPPLP
VRKPGTNDYRPVODLREYNKRYVODIHPVNPVTLNLSALPERKMYTVLIDKAFOL
RLHPTSOPLPEPRMDEGRTGQTLWTRLPQGRKNSPEFDEALHMDLNF10HQH
VTLQYVDLLEAGATKQDLEGTALLLESDIGVRSASKAOICRETVYLGYSR
GGORMLEARKKTVVOIPAPTTAKQVRELFQAFCLMIEGFATLAAPLTYLTKRK
EFSMAPEHQAQFADIKALLSAPALAPVDPKPTLVLDKRRKGAVALYOTLGKRR
PVAYSKLIDPVASGMPVCLKAIAVAALIVKDAKLTGONITVLAHLENTVROP
DRMTNARMTHYOSILLTERVTPAAPPALNPATILPEDEPVTHDHOILLIEETVR
KDIIDIPITGELVLTWTDSSVYVKKMGAAYVDGTRIMASLSPEGSQAKELM
ALTOALRLAEKSTINYDSRYAFATVHAGVATYKQGLITSAGREIKNEELISLE
ALHLEKRLAIIHCPGHQAKADLISRNQMDRVAKQAQAVNLPILEPKAPEPRRO
YTLDEMDKIRKIDQFSETPEGTCTYSGKELIPHEGLLEVVOOIHRLTHGTHLOL
VTSBYHVLRLPGVADSVYKHCVCOLVNNPSRIIPGKRLRGSHGAMHENVTEVK
PAKYGKTLVFDVDFSGWVAEYPTKETSUYAKKLEIEIPREGTPIVYSDNGPA
FVAOYSGIATLIGITDKLHCAYRPOSSGOVERNRTIKETTLTTEGINDMIAL
PVLPRRANTGQGLTPYELLTGPPPLVEIASVHSADVLISQPLSRILKALEWRQ
RAMQOLREAVSEGDLOVPHRFQVGDVSVYRRHAGLWLETRMGKPYLVLLTPTAVK
EGISTWIASHVKLAPPDSDGWRKAEKTENPLRLHLRVYSNNSPGQ"
6112..8085
/gene="env"
6112..8085
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/translation="MHPITLSMRHLPTRGSGPKRLRIPFASIALMTLITTPQSSK
RLIDSSNHPRLSLTWLIDDPGTGVYNSIRGVAPGTWPELHCLRLINAVASTP
PNIYRSYGFCCCTEKEKTCGGSGFSCRMSCVSNISLDDRVKFSFN
SGRQKIVMKLVKDKSCSPSDLYLKISPEKQKQNIQKIMNGSMGIVFKYGGGA
GSTTLIRLETETEPVAPGPKVLEBOGPALEPHNI.PVQULSLRSLSSDTPSY
STGLIPTNTPRNSPGVPVYTGQRLFSLIOGAQOAINSDPATSSCWLCSGGPPY
EGMAKEKFWVTEKHNOCWTGSRNKLTLTEVSGKDCCTGAKPNSHONICNSTVYEO
ASENOYLVGVNMMACNTGLTPCVSTSPFNOSKDCVVOIVPRTYHPEVVIDEY
DYRNRKREPSVLTIAVHLGLSTAVGVTGIDALITGQOLEKSGIETHAATMDLR
ALBESVNLSESLTJSEVVLQNRGLDILFLREGICALEKECCFVYDHSGARIDS
MSEKLEKLERREREADQGEFGMNRSPMTTILSALTGLVLLILITVGPCLIN
RFVAFVERVSAVOIMVLRQYGLISQETDL"
8133..8763
LTR
polyA_signal
8655..8660
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
```

Query Match 50.0%; Score 1659; DB 14; Length 8763;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2259; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 21 agatgcttcttctgctgagatacaaccacttaacacatttgccttcaatg 80
Db 3461 AGATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCTTTTGCTTGAATG 3520
QY 81 gagagatcaagtagaagaacccgagcaccctgaaccgagcttgcctcaaggtt 140
Db 3521 GAGAGATCCAGTAGTAGGGAAGAACCGGCGACCTCACCCTGAGACCCGACCTCCCAAGGCTT 3580
QY 141 caagaactcccgacacatcttgaagaagccctacacaagaagccttgcacactcaagat 200
Db 3581 CAAGAACCTCCCGACATCTTTGACGAAGCCCTACACAGGAGACTGGCCAACTTTCAGAT 3640
QY 201 ccaacacccctcaggtacccctcctcagtaagtgagtgactgctcttgcggaggaacac 260
Db 3641 CCAACACCTCCTCAGGTACCTCCTCAGTAGCTGATGACCTGCTTCTGCGCGGAGCCAC 3700
```

```
QY 261 caaacagagactgcttgaagagtagaaggcactactgcttgaattgtctgactagacta 320
Db 3701 CAAACAGAGACTGCTTGAAGAGTAGAAGGACACTACTGTGAATTGTCTGACCTAGGCTA 3760
QY 321 cagagcctctgtcaagaagagccagatlttgcggagagaggtgaacatacttgggtacag 380
Db 3761 CAGAGCCTCTGCTTAAGAAGGCCAGATTGTGCAGGAGAGAGATMACTACTGGGGTACAG 3820
QY 381 ttgcggagcgagcgacgagatgtgtcagagagacaggaanaactgtatccataacc 440
Db 3821 TTTGCGGGGCGGGCAGCATGTGCGAGGACAGGAGAAACCTGTATGTCAGATVACC 3880
QY 441 ggcgcccaacacagagcccaacagtgagagagttttggggagacgttgaatttgcagact 500
Db 3881 GGCGCCAACACACAGCCCAACAGTGAAGAGATTTTGGGGAGACGCTGATTTTGCAGACT 3940
QY 501 gtgatacccgagggtttgcagacttgcagagcccaacttcaaccgcttaacaaagaaag 560
Db 3941 GTGATACCGGGGGTTTGCAGACTTGAACAGCCCACTTACCGCTTAACCAAGAAAGG 4000
QY 561 gggatctctcgtggtcctcgtgagcaccagagagcatttgaatcacaagaagccctgct 620
Db 4001 GGAATTCCTCTGCTGCTGAGCAGCAGACAGGCAATTTGATGCTATCAAAAGGCGCTGCT 4060
QY 621 gagcgacactgctcgtggccctccctgacgttaactaaaccccttaaccttattgtatga 680
Db 4061 GAGCGCACTGCTGCTGCGCTCTCCCTGAGCTAATCAACCCCTTACCCCTTATGAGATGA 4120
QY 681 gctgaaggagtagccgagagagtttlaaccacaaccctagagacatggagagagactgt 740
Db 4121 GCGTAAGGAGTAGCCCGAGGAGGATTTAACCAACCCCTAGACATGAGAGAGACTGT 4180
QY 741 tgcctactgttcaagaagagctttagctgttagcagatggttggccgtatgtctgaagc 800
Db 4181 TGCCTACTGTGCAAGAAGCTTGAATCCGTGAGCCAGTGCTTGGCCCTATGCTGAAGGC 4240
QY 801 talcgcaactgtgacatactgtgcaagagcgagcgaacaaatttgaacttggagaagaat 860
Db 4241 TATGCCACTGTGGCGCATACTGTGCAAGAGCGCTGAGCAAAATGACTTGGGAAGATAT 4300
QY 861 aactgtlaagcccccaactgtcatgtgagaacatgcttcgagcagcccccaacagatgagat 920
Db 4301 AACTGTATATAGCCCCCATGATGAGAACATGCTTGGCAGCCGCCAAGCAGATGAT 4360
QY 921 gaccaagcccgagctgagcccaacttcaaaagcctgcttctcaagagagagtgcaacttgcg 980
Db 4361 GACCAAGCCCGGACATGACCCACTATCAAAAGCTGCTTCTCAGAGAGGGTCACTTTCGC 4420
QY 981 tccacagcgcctctcaaccctgccaactcttgcctgaagagagacatgaaacagtagac 1040
Db 4421 TCCACAGCGCGCTCTCAACCCCTGCGCACTTCTTGCCTGAAGAGACTGATGAACAGTAGAC 4480
QY 1041 tcatgattgcataactatattgattgagagagactgggttccgaaagagacttacaagat 1100
Db 4481 TCAATGATTGCCATCAACATATTGATTGAGAGAGATGGGGTCCGCAAGACCTTACAGACAT 4540
QY 1101 accgctgacatggagaagtgtaaccttggttcaacttgcggaagagagatgagtgtggaag 1160
Db 4541 ACCGCTGACATGGAAGATGCTTAACCTGTTTCACTGACGGAAGAGCATGATGTGTGGAAG 4600
QY 1161 taagaagatgtctgagcgagcagtggttgaacggaacccgacagatcttggcagcagcct 1220
Db 4601 TAAAGATATGGCTGAGACGCGGTGTGTGAGCGGAGACCCGACAGATGTGGCCACAGACT 4660
QY 1221 ggcggaagagacttgcggaagaaagctgagctcatgtgcccctcaagcaagcttgcgact 1280
Db 4661 GCGGGAAGAGACTTCAAGCGGAAAAGGCTGAGCTCAAGCGCCCTACGCAAGCTTTTGCAGCT 4720
QY 1281 ggcggaaggaatccaataaacttatacggagcaagagatgtgcttggagagagca 1340
Db 4721 GCGCGAAGGAATTCATTAACATTTATACGAGACAGAGATGTGCTTGGAGACTGACA 4780
QY 1341 cgtacaaggggccaatcataaacaaggggttgccttaacctcagcagggaggaataaa 1400
```



		PDKVLAEGPPALPEPNNLPVPLTSLRPDITOPSPNGTGLIPNTPRNSGPVVKI GQRLSLIOGAFOAINSTDPATSSCWLTSSGPPYEGMKEGENTKRNOCITW GSRNKLITLVEAGSGKTGKAPSPSHOHCSTNVYEAKSEMOYLVPGNRMACNTGL TPCVSSVFNOSKDCVWVOIIPRYVYPRVIVLDEHYRFRNRKREVSLTLVMLG LGTAVGVTGTRALITGPQOEKLGELIHAAMTDLRLLESVSLSESLTSLSEVL ONRRLDLFLBEGGLCAKLECCFYVDHSGALRDSMSKLRERLREREREDDOGM FKGMNRPMMWTTLTSLTGPLVLLLTIVGPCIINRFVAFVERSAVQIMVLRQO YQGLSOGERTDL" 2639..6223	
mat_peptide	/gene="pol" /function="retroviral protease and polymerase" /product="Pol protein" 2639..6223		
gene	/gene="pol" /gene="env" /function="retroviral envelope" /product="Env protein" 6099..8069		
mat_peptide	/gene="pol" /gene="env" /function="retroviral envelope" /product="Env protein" 6099..8069		
LTR	8122..8750		
POLYA_signal	8643..8648		
BASE COUNT	2288 a 2214 c 2286 g 1962 t		
ORIGIN			
Query Match 49.3%; Score 1637; DB 14; Length 8750; Best Local Similarity 99.6%; Pred. No. 0; Matches 1987; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY 64	cttttgcttcgaatgtagagatccagatgcggaagaacggcgagctacatgagac 123	QY 724	ccatlggaaggaacactgttgctactctcacaagaagcttgatccctgtagccagtgttg 783
DB 3491	ctttttgcttcgaatgtagagatccagatgcggaagaacggcgagctacatgagac 123	DB 4151	CCATlgGAGGAGaACCTGTTGCTACTACTGTCAAAGAAAGCTTGATCCGTGAGCAGTGGTTGG 4211
QY 124	cgactgcccccaagaggttaagaactcccccgaaccatctttgcaagaagccctcaagaagagac 183	QY 784	cccgatgtctggaagagcatatcgcaagctgtgccaactcgtgcgaagagcgtgcaagaatltg 843
DB 3551	CGACTGCCCCCAAGSGTTTAAAGAACCTCCCGACCATCTTTTGACGAAGCCCTTACACAGGAGAC 3610	DB 4211	CCCGATGCTCGTAAGGCTATGCGACCTGTGGCCCTTACTGGTCAAGGAGCGCTGACCAAAATTG 4270
QY 184	ctgagcaactcgaagatcccaacccctcaagtgagccctcctcagtaagtgtgagactg 243	QY 844	acttggagacagaatataactgttaatagcccccatgtcatgtggaacatcgttcggcag 903
DB 3611	CTGGGCAACTTCAGAGTCAACACCCCTCAGTGACCCCTCTCCAGTGTGATGAGACTG 3670	DB 4271	ACTTTGGAGACAGATATTAAGTGTATAGCCCCCATGTCATTTGGAGAACATCGTTGGCAG 4330
QY 244	cttcgagcggaagcccaacaacagactgcttgaaggtacgaagcaactactgtgaa 303	QY 904	ccccagagacgagatgtgagccaagcccgcatgagccactatcaaaagctgtctcaca 963
DB 3671	CTTCTGGGGGAGCCACCAACAGAGCTGTAGAGGTACGAAGGACACTACTCTGGAA 3730	DB 4331	CCCCAGACCGATGATGAGACCAAGCCCGGATGACCCACATATTAACCTGCTTCACCA 4390
QY 304	ttgtctgagccatgagctacagagcctctgtaagaagccccagatttgcagaagaagagta 363	QY 964	gaagaggtcaacttgcctcacaagccgctcctaacccttcgaactctcgtctggaagag 1023
DB 3731	TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAAGCCCATTTTGCAGAGAGAGGTA 3790	DB 4391	GAGAGGCTCACTTTCGCTCCACAGCCGCTCTCAACCTCTGCTGCTGCTGAGAG 4450
QY 364	acataacttgggtgtagaagtttgcgagcgagcgagctgtagcgagagcgaggaagaa 423	QY 1024	actgtgaacacgtagactatgaltgcatcaactatgtatgtagagagactgggtcgcg 1083
DB 3791	ACATACCTTGGGGTACAGTTTCCGGGGGGGAGCGATGCGTAGAGGAGGACGGAAGAA 3850	DB 4451	ACTGATGAACACAGTACTGATGATGCTATCACTAATTTGATTGAGGAGACTGGGGTCCG 4510
QY 424	actgtatccagatccagcccaacacagcccaaaagtggagaggttlttgggagaa 483	QY 1084	aagagaccttacaagacatccagctgtagcagagagtgtagtaacctgtgtcactgcagagac 1143
DB 3851	ACTGTAGTCCAGATACCGGCCCAACACAGCCCAACAAAGTAGAAGATTGTTGGGACAA 3910	DB 4511	AAGGACCTTACAGACATACCGCTGAGTGAAGATGCTAATCTGTTCACTGACGGAAGC 4570
QY 484	gcttgatlttgcagactgtgatacccgaggtlttgcagacttgaagagcccaacttaccg 543	QY 1144	agctatgtgtgtggaagtagaagatgtcgtgggagcgcaatgtgtgtagcagggagccagc 1203
DB 3911	GCTGATTTTGCAGACTGTGATCCCGAGGGTTTGCAGCTTACAGAGCCCACTACACCG 3970	DB 4571	AGCTATGTGTGGAAGTAGAAGATGTGCTGGGGCGGGTGTGTGAGCGGAGCCACAGC 4630
QY 544	ctaaccaaaagaagaagagatctcctgagctcctgtagcaacagaagaagcatlttgatc 603	QY 1204	atctgggccaagcagctgtccggaagaactcagcgcaaaagctgagctatgagccctc 1263
DB 3971	CTTAACCAAAAGAAAGGGGAATTCCTCTGGGCTCTGAGCACCAAGACATTTGATGATGCT 4030	DB 4631	ATCTGGGCCAGACAGCCCTGCGGAAGGAATCTTACGCCCAAAAGGCTGAGCTCATGGCCCTC 4690
QY 604	atcaaaaagccctgtgagcgacactgtcttgccctcctgtagcagtaactaaacctt 663	QY 1264	acgcaagcttgcgagcgtgcggaagaagaatccataaacttatacgaagcaggtat 1323
DB 4031	ATCAAAAAGGCCCTGCTGAGAGCACTGCTCTGAGCTCTCTGAGCTGATTAACCTTTT 4090	DB 4691	ACGCAAGCTTGTGGGCTGCGCGGCGGAAGGAATTCATTAACCTTATACGGGACACAGGTAT 4750
QY 664	acccttatagtgtagcgtlaagagtagcccgagaggtttaaaccacaacccagta 723	QY 1324	gaccttgcagctgcacagctacacgagggccatcatalaaacaagggggtgtgctactca 1383
DB 4091	ACCCTTATGTGATGAGCGTAAGGAGGTAGCCGAGAGATTTTAACCCAAACCTTAGGA 4150	DB 4751	GCCTTGGCAGCTCACACAGTACACGGGGCCATCTTAAACAAAGGGGGTGTGCTTACTCTCA 4810
QY 1384	gcagggagggaaataaagaacaaagaggaattcgaacctatagaagccctacattg 1443	QY 1444	ccaaaaggtcagctatctatcaactctcctgagacatcagaagaacaaagatcctatct 1503
DB 4811	GCAGGGAGGGAAATTAAGAAACAAGAGAAATTTTAAGCTTATTAAGACCTTACATTTG 4870	DB 4871	CAAAAAGGCTAGCTTATTAACACTCTCTGGAACATCAAAAGCCCAAGATCTCATATCT 4930
QY 1504	agaaggaacacagatgtgctgacgggtgtgccaagcagagcagccagagctgtttaaactctg 1563	QY 1504	agaaggaacacagatgtgctgacgggtgtgccaagcagagcagccagagctgtttaaactctg 1563
DB 4931	AGAGGGAACACAGATGCTGACCGGTTGCAAGCAGGAGCCAGGCGCTTTAAACCTTCTG 4990	DB 4931	AGAGGGAACACAGATGCTGACCGGTTGCAAGCAGGAGCCAGGCGCTTTAAACCTTCTG 4990
QY 1564	cctatatagaagacggcccaagccccaagaaacccaagacagcagtaacccctagaagactg 1623	QY 1564	cctatatagaagacggcccaagccccaagaaacccaagacagcagtaacccctagaagactg 1623
DB 4991	CCTATATATGAAGAGCCCAAGGCCCAAGCCCAAGCAGCAGCACTATACCTTAAGAACTGG 5050	DB 4991	CCTATATATGAAGAGCCCAAGGCCCAAGCCCAAGCAGCAGCACTATACCTTAAGAACTGG 5050
QY 1624	caagagataaaagaataagacaggtctctgagaccccgagagggagacgtgtcctatccca 1683	QY 1624	caagagataaaagaataagacaggtctctgagaccccgagagggagacgtgtcctatccca 1683
DB 5051	CAAGAGATTAAGAAAGATTAACCAAGTCTCTGAGACTCCGAGGGGAGACCTGCTATACCTCA 5110	DB 5051	CAAGAGATTAAGAAAGATTAACCAAGTCTCTGAGACTCCGAGGGGAGACCTGCTATACCTCA 5110
QY 1684	tatgggaagaaatccctgcccccaagaagaaggtttagaatalgtlccacaagatacatcgt 1743	QY 1684	tatgggaagaaatccctgcccccaagaagaaggtttagaatalgtlccacaagatacatcgt 1743
DB 5111	TATGGGAAGGAATCTCTGCCCAACAAGAGGTTATGAATATGTCCAAAGATACATGCT 5170	DB 5111	TATGGGAAGGAATCTCTGCCCAACAAGAGGTTATGAATATGTCCAAAGATACATGCT 5170
QY 1744	ctaacccacttgaagacttaaacacactgcagagaggtgtgtcagaacatccctatatacgt 1803	QY 1744	ctaacccacttgaagacttaaacacactgcagagaggtgtgtcagaacatccctatatacgt 1803
DB 5171	CTCACCCACTTAGGAATTAACCACTGACGAGCTGTGTGTCAGAACTCCCTTATCATGTT 5230	DB 5171	CTCACCCACTTAGGAATTAACCACTGACGAGCTGTGTGTCAGAACTCCCTTATCATGTT 5230



QY 331 gctaaagaagccagaattgcagagagaagtaacatacttgggttacagtcttcgagc 390  
DB 3847 GCTAAAGAGCCAGATTGGAGAGAGTAACTACTTGGGTACAGTTTGGGGGCG 3906  
QY 391 ggcagcagatgctgaacagagcacgagaacaaactgagtcacataccggcccaac 450  
DB 3907 GGGCAGCAGTGGCTGACGAGGACGAGAGAGAACTGATGCAATACCGGCCCAAC 3966  
QY 451 acagcaaaacagtgagagaatttgggagacgttgatttgacagcttgatcccg 510  
DB 3967 ACAGCAAAAGTAGAGAGTTTGGGAGAGCTGATTTGGAGCTGTGATCCG 4026  
QY 511 gggttggagactttagcagcccaactctaccgctaaccaagaagaagggaattcc 570  
DB 4027 GGGTTGGCAGCTTAGCAGCCCACTTACCCACTAACCAAGAAAAAGGGGAATTC 4086  
QY 571 tgggctcctgagcacagagaagcatttgatctatcaaaaggccctgctgaagcact 630  
DB 4087 TGGGCTCCTGACACCAAGAGCATTTGATGCTATCAAAAAGGCCCTGAGGCACT 4146  
QY 631 gctctggccctccctgaacgtaaaccttacccttlatgtgatgagcgtlaagga 690  
DB 4147 GCTCTGGCCCTCCCTGAGGTGACTTAACCTTTACCTTTATGTGATGAGCCTAAG 4206  
QY 691 gtagcccaaggagttttaaccacaacccctagagacaaaggaagagacctgtgc 750  
DB 4207 GTAGCCCGGGGAGTTTAAACCAACCTTAGACCATGAGAGAACCTGTGTGCTAC 4266  
QY 751 tcaaaagagctgatacctcgttagcagagtggttgcgcgtlatctggaaggtatcc 810  
DB 4267 TCAAAAGAGCTGATCTGTAGCCAGTGTGTGGCCCATATGCCGTAAGGCTATG 4326  
QY 811 gtagcgaactggtgtaagaagcgtgacaaattgacttggagacagaaataactgt 870  
DB 4327 GTGGCATACTGCTCAAGAGAGCTGACAAATTGACTTTGGGACAGAAATTAAC 4386  
QY 871 gcccacatgcatltaggaacatcgttcgagagcccccagacagatgatacgaagc 930  
DB 4387 GCCCCCATGCGTTGGAGAAACATGTTGGGAGGCCCAACCGATGAGTACCAAC 4446  
QY 931 cgcataccacatacaaaagcctgtctctacagagaaggttcaacttgcgtccacagc 990  
DB 4447 CGCATAGACCCACTATCAAAAGCTGCTCTCAAGAGGGTCAACGTTGCTCCACAG 4506  
QY 991 gcttcaaaccttgcacatcttctgcctggaagagactgataaacagtlactatga 1050  
DB 4507 GCTCAACCTCGCACCTCTCTGCTGGAAGAGACTGATGAACCAAGTACTCATG 4566  
QY 1051 catcaactatgatttggagagactgggttcgcaagaagccttcaacataccgctgact 1110  
DB 4567 CATCAACTATTGATTGAGAGAGACTGGGGTCCGCAAGAGACTTACAGACATAC 4626  
QY 1111 ggaagaagtgtacacctggttactgaagagagactatgtgtggaagtagaagagt 1170  
DB 4627 GGAAGAAATGTTAACTGTTACACTGAGAGCAAGCATATATGTGAGAGTAA 4686  
QY 1171 gctgggagcagtggtgtgagcgggagcccgacagatctgggcaagcagcctgcgga 1230  
DB 4687 GCTGGGGGGGGGTGGAGAGCGAGCCGACAGCATCTGGGCCAGCGCTCCGGA 4746  
QY 1231 acttcaagcgaagaagctgagctcaatggccctcaagcaagcgttggggtgagcga 1290  
DB 4747 ACTTCAACACAAAAGGTGAGCTCATGCGCTTACCAAGACCTTTGGGGTGG 4806  
QY 1291 aaatcaataacattatagcagacagagtgatgcttgcagctgacagctagacgg 1350  
DB 4807 AAATCCATAAACAATTAACAGACAGAGTATGCTTGGCAGCTGCAACCTAC 4866  
QY 1351 ggcataataaagaagggttgccttactacagcagggaggaagaataaagaagaag 1410  
DB 4867 GGCATCTATAGCAAAAGGGGTGCTTACCTCAAGAGGGAGGAATAAAGAAAG 4926  
QY 1411 gaaattcaagcctattagaagccttacaattggcaaaagcctagctattatagcgt 1470  
DB 4927 GAATTTCTAAGCCTATTAGAGCCTTACATTTGCCAAAAAGCTGACTATTA 4986  
QY 1471 cctgagacatcagaagccaaagatctacataltatagaggaacagatgctgacagg 1530  
DB 4987 CCTGAGACATCAGAAAGCCAAAGATCCATATCCAGAGGGAACCAATGGCTGAC 5046  
QY 1531 gccaagcagagcccaagcgttgaactcttgcctataatagaagcccaagccca 1590  
DB 5047 GCCAAGCAGCAGCCAGGGGTGTAACCTTGCCTATGATAGAAACCCCAAGCC 5106  
QY 1591 gaaccagacagcaatagacccctagaagactggcgaagagataaagaatgaccagtt 1650  
DB 5107 GAACCGGAGCAGACATGACACCTTAGAAGACTGCGCAAGAAATTAAGATGAC 5166  
QY 1651 tctgagactcggagaggagcctgtatacctcataatgggaagaaatccctccaca 1710  
DB 5167 TCTGAACCTCCGAGAGGACCTGTATACCTCAGATGGGAAGAAATCTGCCCCA 5226  
QY 1711 gaagggttagaatgttccaaagatatacgtctaaaccactagaactaaacactg 1770  
DB 5227 GAAGGTTTAGAATATGTCACAGATACATGCTTAACCCACTAGGAACCTAAC 5286  
QY 1771 cagcagttgltcagaacalcccttatacatgttctgaagctaccagagtgagctg 1830  
DB 5287 CAGCAGTTGTCAGAAATCCCTTATCATGTTCTGAGAGCTACCAAGAGTGGCTG 5346  
QY 1831 gtagtcaaacatgtgtgctcctgcagcgttgatgaactcctccagaatacctca 1890  
DB 5347 GTGTCAAAACATGTGTGCTCCCTGACAGCTGTGTATGCTAATCCTCAAG 5406  
QY 1891 ggaagagacataaaggagagccacacagcgtcactacggaggaaggttcactggt 1950  
DB 5407 GGAAGAGACTAAGGGAGAGCCACCCAGGCGCTACTAGGGAGGAGCTTCACTG 5466  
QY 1951 aagcgcgtataaagaaacaaalataatgtgttcttgaacaccttcaagagt 2010  
DB 5467 AAGCGGCTTAATAGCAAAACAAATATCTATGTTTTTGAAGACACTTTAG 5526  
QY 2011 gtagaggttacttactaagaagaagactcaacggtgtgtgtaagaataactgag 2070  
DB 5527 GTAGAGGCTTATCTTACTTAAGAAAGACATCAACCTGTGTGCTAAGAAAT 5586  
QY 2071 gaaatttccaagatttgaatlaacctaaggtlaaggtcagaatgltcagctt 2130  
DB 5587 GAATTTTTCCAAATTTGCAATTTGCAATACCTAAGGTATAGGCTCAACATG 5646  
QY 2131 gtagcccaagtgatgaagcagctgccaagataatgggagattgattggaactgact 2190  
DB 5647 GTTCCAGGTAAAGTCAAGGAGCTGGCCAAAGATATTGGGATGATTGA 5706  
QY 2191 gcatcagagcccaagcctcagaacagtgtagaagaatgaataacatlaaagacc 2250  
DB 5707 GCATACAGACCCCAAACTCAGACAGTGAAGAGATGAATGAACCATTAAGAG 5766  
QY 2251 cttaaccaatltgaccacagagactgcatlaagattgattgagctctcctgcctt 2310  
DB 5767 CTTAACCAATTTGACCAAGAGACTGGCATTAATGATTGATGCTTCTGCCCTT 5826  
QY 2311 cttttaggttgaagaaacacccctggaagatttggcgtgacccctatgaattgtct 2370  
DB 5827 CTTTATTTAGGGTGAAGAACCCCTGAGACAGTTTGGGCTGACCCCTATGA 5886  
QY 2371 gggggaaccccccgttggagaagaatttgccttgcacatagtgctgacttcc 2430  
DB 5887 GGGGAGACCCCCCGTTGGCAGAAATGCTTGGCAATAGTGCTGATGTGCTTCC 5946  
QY 2431 cagccttgttctcagctcagaagcgtcagatggttgaagcagagcgtggaagag 2490  
DB 5947 CAGCCTTTGTTCTTCTAGGCTCAAGGCGCTCGAGTGGGTGAGGACGAGCG 6006  
QY 2491 ctccggagagcctactcagaggaagagacttgcaggtccacatcgcttccaggtt 2550



Db 6007 CTCGGGAGGCGCTACTCAGGAGAGATTGCAATGTTCCATCGCTTCCAACTTGGACAT 6066  
 Qy 2551 tgaattatgttagaagccacgctgcaagaaacctcgaactcgg tggaaagggacctat 2610  
 Db 6067 TCAGCTGATAGTAAAGCCGACGCTGAGGAAACCTCGAGACTCGGTGGAAGGACCTTAT 6126  
 Qy 2611 ctgcgtactttgaacacacacgctgtgaaagtcgaagaaatcccaactcgtgctcat 2670  
 Db 6127 TCCTGCTTGTGACACACACGAGGCTGTGAAGATGCAAGGAAATCCCACTGATCCAT 6186  
 Qy 2671 gcatccacgcttaagccgacgacacccctcgaactcgaaggtggaagacgaaagactga 2730  
 Db 6187 GCATCCGCTTAAGCCGCGACCTCCGACTCGGGGTGGAAGGCAAAAGACTGAA 6246  
 Qy 2731 aatcccttaagcttcgactcgaactcgaactcgaactcgaactcgaactcgaactcga 2790  
 Db 6247 AATCCCTTAAGCTTGCCTCCATCGCTGCTGCTTACTGCTCATATACCTCCTCACT 6306  
 Qy 2791 taatgttaaacgctgtgtgaaagccggaactcccaataacctatctcactcgtgt 2850  
 Db 6307 TAATGCTAAACGCTTGTGACACGCGGAACTCCCAATACCTTATCTCTCACTCGTT 6366  
 Qy 2851 actactgactccgctgacagcttaataataacagcactcaagggagggcctccttgg 2910  
 Db 6367 ACTTACTGACTCGGCTGACAGGATTAATTAACAGCAGCTCAAGGGAGGCTCCCTTGGG 6426  
 Qy 2911 gacctgtgctgactgaattatgtctgctcctgcatgaatccctcgtgtcctaatagca 2970  
 Db 6427 GACCTGTGCTGCTGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6486  
 Qy 2971 ggcac 3030  
 Db 6487 GCGCCACACCCCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6546  
 Qy 3031 taatgaagaatattgtgaaatccctcgaagctccctcgaagcaatgagctcgaacta 3090  
 Db 6547 TAATGAAGAATATTGTGGAATCCTCAGATTTCTTTTGCACCAATGAGGCTGCAATAC 6606  
 Qy 3091 ttctaagatggaatgtgaatggaatggaatggaatggaatggaatggaatggaatggaat 3150  
 Db 6607 TTCTAATGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 6666  
 Qy 3151 ttttaacaatcctacacagcttataatcaatlaattatgacatggaatggaatggaatgga 3210  
 Db 6667 TGTAAACATTCCTACAGTATTAATCAATTAATTAATTAATTAATTAATTAATTAATTA 6726  
 Qy 3211 gcaacacgaggtacacaaagatgacaaagaaataagcaataagctgcatcgtttagact 3270  
 Db 6727 GCAACACGCGGTACAAAGATGTACGAATAAGCAATAAGCTGTCATTCGTTAGACT 6786  
 Qy 3271 agactacttaaaataagtttact 3295  
 Db 6787 AGATTACTTAATAAATAAGTTTCACT 6811

RESULT 10  
 AF038601 7333 bp mRNA MM 21-MAY-1998  
 LOCUS AF038601  
 DEFINITION Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete  
 sequence.  
 ACCESSION AF038601  
 VERSION AF038601.1 GI:3133304  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 REFERENCE  
 1 (bases 1 to 7333)  
 Fishman, J.A., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P., and  
 Fishman, J.A.  
 Identification of a full-length cDNA for an endogenous retrovirus  
 of miniature swine  
 J. Virol. 72 (5), 4503-4507 (1998)

MEDLINE  
 98216827  
 REFERENCE  
 2 (bases 1 to 7333)  
 Fishman, J.A.  
 AUTHORS  
 Direct Submission  
 TITLE  
 Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts  
 JOURNAL  
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REFERENCE  
 3 (bases 1 to 7333)  
 Fishman, J.A.  
 AUTHORS  
 Direct Submission  
 TITLE  
 Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts  
 JOURNAL  
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REMARK  
 Sequence update by submitter  
 COMMENT  
 On May 16, 1998 this sequence version replaced gi:3116445.  
 FEATURES  
 source  
 1..7333  
 /organism="Sus scrofa"  
 /strain="miniature swine"  
 /db\_xref="taxon:9823"  
 /cell\_line="PK-15"  
 /note="porcine endogenous retrovirus ERV-PK15, infective  
 virus but represents a naturally occurring transcript"  
 1..578  
 598..2172  
 /codon\_start=1  
 /product="gag protein"  
 /protein\_id="AAC1676.1"  
 /db\_xref="GI:3116446"  
 /translation="MGQTVTPPLSLTLDHWEVRSRAHNLVQVKKGPWQFCASEWP  
 TFDGMSSEGFENSEILAVKALIFOTGPGSHDOEPLYLTMODLADPPWPKMLN  
 KPRKPGRIIALGKRNHSAKEVSPRIYELIEPTWEPQPPPPYPAQGAARG  
 PSAPPGAIVYEGPACGTRSRGATBERTDAILPLRTYGPMPGGLOPLOYWFS  
 ADLTMKTNHPPESDEQRLGLVESLMPSHQMDCCOOLQTLFTTERERILLEA  
 KKNVGDAGRTQDNEIDMFLTRGMDYNTABEGRESKITRQALVGLRASRRP  
 TNLAKREVMQGPNEPVSFLERMEAFRRTPDPYSEAKASVALATIGGSAIDIR  
 KILQRLGLQEAELRDLVREAEKYVRRTEEEKEREREREERERRRORCKNLT  
 KILAAVEGKSRERERDFKIRSGPQSGNLGNPFLDKDQACAYCEKKGMAHNCPEK  
 KGNKGPVIALIEDKD"  
 2320..3552  
 /note="truncated protein"  
 /codon\_start=1  
 /product="pol protein"  
 /protein\_id="AAC1676.1"  
 /db\_xref="GI:3133305"  
 /translation="MGATGQROYPMTRRVYDLDGVHVSFLVIEPCPYLLGRDLL  
 TKMGATSEFGREPREVSNKPIVTLTDLDDERLSPQVKKPPRODIOSLMEQFQAM  
 AETAGMGLAKVPPQVYDLSKATPVSVROYVRSREARIGMPVORLIQOGLIVPYO  
 SPWNPFLVLRKPGTNDIRPVQDILREVKRKQDILHPVPPVNLSSALPBERMMYVL  
 DLKDAFELRLHPQSQPLETEWMDPGTGRGQLMTRLPDGFKNSTPTTDEALHRL  
 ANFRIOHPOVTLQYVDLLLAGATKQDCLGTRALLLEISDLYRASAKAQICRE  
 VTYLGYSLRGQRMILTEARKKTVVQIPAPTYAKOVREFIAGFCRLMIPGFATLAP  
 LYPLTKKGGCLPQGGK"  
 4803..6725  
 /note="env protein: N terminus deleted"  
 /codon\_start=1  
 /pseudo  
 6726..7333  
 LTR  
 BASE COUNT 1984 a 1788 c 1885 g 1676 t  
 ORIGIN  
 Query Match 32.9%; Score 1091; DB 4; Length 7333;  
 Best local Similarity 99.7%; Pred. No. 0;  
 Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1364 aaaggagggtgtgcttacctgaagcagggaggaataaagaacagaggaatctcaagcc 1423  
 Db 3517 AAAGGGGGTGTCTTACTCAGCGAGGAAATTAAGAAAGGAAATTTTAAAGCC 3576  
 Qy 1424 tattagaagccttaccattgccaagaagctagctatatacactgtccctggacatcaga 1483  
 Db 3577 TATTAGAGCCTTACATTTGCCAAAAGCTACTCTATTATACGCTGCGGACATCAGA 3636  
 Qy 1484 aagccaaagatctcatatctagaaggaaacagatggtcagcggtgtgccaagcagcgag 1543



Db 3637 AAGCCAAAGATCTCATATATTAAGGAGAACCAATGGCTGACCGGGGTTGGCAAGCAGGAG 3696  
Oy 1544 cccagcgtgtaacctctgctataatagaacgcccagcccccagaaccagaagac 1603  
Db 3697 CCCAGGCTTTAACTTCTGCTCTATATATGAACCCCAAGCCCAAGACCCCAAGACGAC 3756  
Oy 1604 agtaacacctagaagactggaagaagataaaaaagatagaccagttctctgagactcgy 1663  
Db 3757 AGTACACCTTAGAAGACTGGCAAGAGATAAAAAGATAGACAGTTCTCTGAGACTCCGG 3816  
Oy 1664 aagggagcctgtatactataatggaagaagaaacctgcccacaagaagaggttagaat 1723  
Db 3817 AAGGGAGCTGTATACCTCATATATGGAAGAAATCCTGCCCAAAAGAGGTTAGAT 3876  
Oy 1724 atgtccaacagataacgctcgaacccaactaggaactaaacactggaacgattgagca 1783  
Db 3877 ATGTCCAAACAGATACATGCTTAACCCACCTAGGAACCTTAACACCTGCAACGAGTTGCTCA 3936  
Oy 1784 gaacatccctcatcatgltctgaggtcaccagagtgctgactcgytggtcaacaat 1843  
Db 3937 GAACATCCCTTATCATGTTCTGAGAGCTACGAGAGTGGCTGACTCGGTGCTCAACAT 3996  
Oy 1844 gtgtgcccctgccaagctgtgtaagtatacctccagaatacctccaggaagaagataa 1903  
Db 3997 GTGTGCCCTGCGACGCTGTTATGCTAATCCTCCAGATACCTCCAGAAAGAGACTAA 4056  
Oy 1904 ggggaagccaccagcgcctacgtggaagtggacttcaagtgaagtaagccgctaata 1963  
Db 4057 GGGGAAGCCACCAGGCGCTCCTGGAAGTGGACTTCACTGAGGTTAAAGCGCGCTAAT 4116  
Oy 1964 acggaacaataatcatatgtgttctgtagaaccttccagatgggttagaggtatc 2023  
Db 4117 ACGGAACAATAATCTATGTTGTTTGTAGACACCTTTTCAGAGTGGTTAAGGCTTATC 4176  
Oy 2024 ctacttaagaagaagactccaacgtgtgtgctagaanaaatacctggaagaaatlttccaa 2083  
Db 4177 CTACTTAAAAAAGACTTCAACCGGTGGCTAGAGAAATCTGGAGGAAATTTTTCCAA 4236  
Oy 2084 gattgggaatacctaaggttaaggttagaagaacatggtccagcgttctgttccaggtaa 2143  
Db 4237 GATTGGGAATACCTTAAGGTAATAGGGTGAACATGCTCCAGCTTTCCTTCCCAAGTAA 4296  
Oy 2144 gtcaggaactggaagaatattggygatttga tggaaactgcatgtgcatatacagaccc 2203  
Db 4297 GTCAAGGACTGGCCAAAGATATTGGGGATTGATTGAAACCTCATTTGTCATACAGACCC 4356  
Oy 2204 aaagctcagaagtagaagaagatgaataaacatlaaagaagacccttaaccaatga 2263  
Db 4357 AAAGCTCAGACAGGTAGAGAGATGAATGAACCATTTAAAGAGAACCTTACCAAAATTGA 4416  
Oy 2264 ccacagaactggaatgatattgagatgctccgacctgtgttctttagaggtga 2323  
Db 4417 CCACAGAGACTGCGATTATATGATGGATGCTCTCCGCTTGTGCTTTTAAAGGATGA 4476  
Oy 2324 ggaacacccctggaacagtttgggctgaacccctta tgaattgcttaagcggggagacccccc 2383  
Db 4477 GGAACACCCCTGAGACAGTTTGGGCTGACCCCTTAATAATTGCTCTACGGGGAGACCCCCC 4536  
Oy 2384 cgttggcagaagaattgaccttgacatatgctgtagtgcgtgcttcccaagccttgttct 2443  
Db 4537 CGTTGGCAGAAATTGCTTTGACATATGCTGATGTGCTGCTTTCCACAGCTTTGTCT 4596  
Oy 2444 cttaggctaaagcgtcgaagtgggttgaagcagaagcgttgaagcagctccggaagcct 2503  
Db 4597 CTAGGCTCAAGGCGCTGAGTGGGTGAGGCGACGAGCTGGAAGCACTCCGGAGGCGCT 4656  
Oy 2504 actcagaaggaagacttgcaagttccacatcgcttccaagttggaaltcagttcatgta 2563  
Db 4657 ACTCAGGAGGAGACTTGGCAAGTTCCACATGCTTCCAAAGTTGAGATTCAATCATGTA 4716  
Oy 2564 gaccacacgtgtagaagaacctggaagctggtggaaggaacctatctgtaactttga 2623

Db 4717 GACGCCACCGTGCAAGAAACCTCGAGACTCGTAGAAGGAGCACTTATCTGACTTTTGA 4776  
Oy 2624 ccacacacagcgtgtgaagtcgaagaatcccc 2658  
Db 4777 CCACACCAAGCGCTGTGTAAGTTCGAGAGAAATCCCC 4811

RESULT 11  
AR130474  
LOCUS AR130474 7333 bp DNA  
DEFINITION Sequence 2 from patent US 6190861.  
ACCESSION AR130474  
VERSION AR130474.1 GI:14118799  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7333)  
AUTHORS Fishman, J.A.  
TITLE Molecular sequences of swine retroviruses method of using  
JOURNAL Patent: US 6190861-A 2 20-FEB-2001;  
FEATURES  
source location/Qualifiers  
I. 7333  
BASE COUNT 1984 a 1788 c 1885 g 1676 t  
ORIGIN

Query Match 32.9%; Score 1091; DB 6; Length 7333;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1364 aaaggggtgtgttactctcgcagcggggaagaaataaagaacaaaggaagaaatcttaagc 1423  
Db 3517 AAAGGGGCTTGTCTTACTCAGCAGGAGGGAATAAAGAAACAAGAGAAATTTCTAAGCC 3576  
Oy 1424 tattaagaaccttaacattccaaagagctagctattatatacctgtccctgagcatcaga 1483  
Db 3577 TATTAGAAGCTTATCATTTTGCAGAAAGGCTAGCTATTATACCTGTCTCGGACATCAGA 3636  
Oy 1484 aagccaaagatcctcatatctagagggaaaccagatgctgacgggttgcgaagcagcg 1543  
Db 3637 AAGCCAAAGATCTCATATCTAGAGGAGCAACAGATGCTGACGGGTTGCCAAGCAGGAG 3696  
Oy 1544 ccaaggtgttaacctctgtcctataatagaacgcccagaagcccaagaaccagaagcgc 1603  
Db 3697 CCAAGGCTGTAACTTCTGCTTATATGAAGAGCCCAAGGCCAGAACTCCAGACGAC 3756  
Oy 1604 agtaacacctagaagactggaagaagataaaaaagatagaccagttctctgagactcgy 1663  
Db 3757 AGTACACCTTAGAAGACTGGCAAGAGATGAATGAACCATTTAAAGAGAACCTTACCAAAATTGA 4416  
Oy 1664 aagggagcctgtctatacctatattggaagaagaaatcctgcccacaagaagaggttagaat 1723  
Db 3817 AAGGGAGCTGTATACCTCATATATGGAAGAAATCCTGCCCAAAAGAGGTTAGAT 3876  
Oy 1724 atgtccaacagataacgctcgaacccaactaggaactaaacactggaacgattgagca 1783  
Db 3877 ATGTCCAAACAGATACATGCTTAACCCACCTAGGAACCTTAACACCTGCAACGAGTTGCTCA 3936  
Oy 1784 gaacatccctcatcatgltctgaggtcaccagagtgctgactcgytggtcaacaat 1843  
Db 3937 GAACATCCCTTATCATGTTCTGAGAGCTACGAGAGTGGCTGACTCGGTGCTCAACAT 3996  
Oy 1844 gtgtgcccctgccaagctgtgtaagtatacctccagaatacctccaggaagaagactaa 1903  
Db 3997 GTGTGCCCTGCGACGCTGTTATGCTAATCCTCCAGATACCTCCAGAAAGAGACTAA 4056  
Oy 1904 ggggaagccaccagcgcctacgtggaagtggacttcaagtgaagtaagccgctaata 1963  
Db 4057 GGGGAAGCCACCAGGCGCTCCTGGAAGTGGACTTCACTGAGGTTAAAGCGCGCTAAT 4116  
Oy 1964 acggaacaataatcatatgtgttctgtagaaccttccagatgggttagaggtatc 2023

Db 4117 ACGGAACAAATATCTATGTTTGTAGACACCTTTGACGATGGGTAGAGCTTATC 4176  
 Oy 2024 ctactaagaagaagacttcaacgcttgctgctgaagaataactggaagaattttccaa 2083  
 Db 4177 CTACTAAAAAGAACTTCAACCGTGTGGCTAAGAAAAATACGTGAGAAATTTTTCAA 4236  
 Oy 2084 gatttgaataacctaaagtaagatgacagaactgtgccagcttctgcttgcaggttaa 2143  
 Db 4237 GATTGGAAATACCTAAGGTATAGGGTCAGACAAATGGTCCAGCTTTGGTGGCCAGGTAA 4296  
 Oy 2144 gtccagagacttgcgaagataattggtgattggaacttgcatttgcatacagaaccc 2203  
 Db 4297 GTGAGGAGCTGGCCAGATATTGGGAGATTGATAAACTGCATTGTGCATACAGACCCC 4356  
 Oy 2204 aaagctcagaagctgtagaagagatgataaaccattaaagaaccccttaaccaattga 2263  
 Db 4357 AAAGCTCAGACAGTATAGAGATGATGATGAACCATTTAAAGAGACCCCTTACAAATTTGA 4416  
 Oy 2264 ccacagaagactgcatlaatatgattgagctcctccctgcttctttagagggtga 2323  
 Db 4417 CCACAGAGACTGGCATTTATGATTGAGATGGCTTCTGCTGCTTTGTGCTTTTAGGGTGA 4476  
 Oy 2324 ggaacaccccttgagacagtttggctgaccccttatgaaattgcttaacggggagacccccc 2383  
 Db 4477 GGAACACCCCTGAGACAGTTTGGGCTGACCCCTTAAATTTGCTTACGGGGGACCCCCC 4536  
 Oy 2384 cgttggcagaagaattgcttgcacatagtgctgagtgctgcttccagaccttcttct 2443  
 Db 4537 CGTTGGCAGAAATTTGCTTTCACATAGTGTGCTGATGCTGCTTTCCAGCCTTTGTTCT 4596  
 Oy 2444 cttagctcaagcgctcagctgagtggttgaagcagcagcgttgaagcagctcccgaggagcct 2503  
 Db 4597 CTAGGCTCAAGGCGCTGAGTGGGTGAGAGCAGCAGCGCTGGAACAGACTCCGGAGGACCT 4656  
 Oy 2504 actcagggagagacttgaagttccacatgcttccaaagttggagattcagttatgtta 2553  
 Db 4657 ACTCAGGAGAGACTTGGAACTTGCACATTCCTTCCAGTTGGAGATTCACTATATGTA 4716  
 Oy 2564 gacgcacacgctgcaagaaactcagactcgttgaagagagacattatctgacttga 2623  
 Db 4717 GACGCCACCGTGCAGAGAACTCCAGACTCGGTAGAAAGGACGCTTATCTGCTACTTTGA 4776  
 Oy 2624 ccacacaaacgctgtgaaagtcgaagaatcccc 2658  
 Db 4777 CCACACCAACGCTGTGAAGTCGAAGGAATCCCC 4811  
 RESULT 12  
 AX052635 4402 bp DNA PAT 12-JAN-2001  
 LOCUS AX052635  
 DEFINITION Sequence 20 from Patent WO0071726.  
 ACCESSION AX052635  
 VERSION AX052635.1 GI:12226825  
 KEYWORDS  
 ORGANISM porcine endogenous retrovirus.  
 SOURCE porcine endogenous retrovirus  
 RETROVIRUSES: Retroviridae; Mammalian type C  
 TITLE 1 (bases 1 to 4402)  
 JOURNAL Federpietel,M.J.  
 PATENT: WO 0071726-A 20 30-NOV-2000;  
 xenotransplantation  
 Patent: WO 0071726-A 20 30-NOV-2000;  
 Mayo Medical Ventures (US)  
 FEATURES  
 SOURCE Location/Qualifiers  
 1..4402  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 1259 a 1085 c 1111 g 947 t  
 ORIGIN

Query Match 25.5%; Score 848; DB 6; Length 4402;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 2198; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
 Oy 31 tctctcctagatttaacccacatagccaaacatttctgcttctggaatggagatcca 90  
 Db 1043 TTCTGCTTGAATTTACACCCACCTACCAACACATTTTGGCTTGAATGAGATCA 1102  
 Oy 91 gttacggagaagaacggcgagctcacttggacccgactgcccagaaggttcaagaactcc 150  
 Db 1103 GGTACGGGAAGAAACCGGACGCTCACCTTGACCACCTCCCAAGGTTCAAGAACTCC 1162  
 Oy 151 ccgacacatttgcagaagccctacacagaagacttgcgaacttaagataccaacct 210  
 Db 1163 CCGACCATCTTTAGCAGAACCCCTACACAGACCTGGCCCACTTGAAGATCCAAACCTT 1222  
 Oy 211 caagtgaacctctctcagtagcgttggatgaccttcttggggagaccccaagaagac 270  
 Db 1223 CAGGTGACCTCTCCAGTACGATGATGACCTGCTTCTGGGGAGCCACCAACAGAC 1282  
 Oy 271 tgcctagaaggtacagaagcaactactgcttgaatttctgacctaggtctacagacctct 330  
 Db 1283 TGCTTTGAAGGACGACGAAGGACTACTGCTGATTTGCTGACCTAGGCTACAGAGCTCT 1342  
 Oy 331 gctaaagaagccagatttgcaggaagaaggttaacatatttgggttaacagtttgcggggc 390  
 Db 1343 GCTTAAGAGGCCCCAGATTGTCAGAGAGAGATTAACCTTGTGGGTACAGTTTGGGGGAC 1402  
 Oy 391 gggcagcagatgcttgcagcagagcaggaagaacagttagtcagaataccagcccaacc 450  
 Db 1403 GGGCAGCGATGCTGACGAGGACGACGGAAGAAACTGTGTCAGATACCGGCCCAAC 1462  
 Oy 451 acagccaacaagltgaagagtttctgggacagcgttgaatttgcagacttgcagctccg 510  
 Db 1463 ACAGCCAACAAGTGAAGAGAGTTTGGGGACAGCTGGAATTTGACAGATGTGATCCCG 1522  
 Oy 511 ggggttgcagccttaagcagcccaactcctaccgttaaccaagaagaaggagattctcc 570  
 Db 1523 GGGTTTGGCAGCTTACAGGCCCTCACCTCTACCCCTAACCAAAAGAGGGAATTTCTCC 1582  
 Oy 571 tgggtctccttaagcagcaggaagcatttgaatgctatacaaaagcccttgcagcact 630  
 Db 1583 TGGGCTCTGACGACCAAGAAAGCATTTGATGCTATMAAAAGGCCCTGCTGAGCGACCT 1642  
 Oy 631 gctctggccctccctgacgttaactaaacctttacccttatagttgaatgaagcgtlaagg 690  
 Db 1643 GCTCTGGCCCTCCCTGACGTAACTAAACCTTTTATGATGATGAGCGTAAGGGA 1702  
 Oy 691 gtagcccgagagattttaaaccacacccatagagccatgtagagagacttctgctaccgt 750  
 Db 1703 GTAGCCCGGGGAGTTTAAACCAACTTGAAGCAATGAGAGACCTGTGTGCTACCTG 1762  
 Oy 751 tcaagaagcttgatccctgagcagtggttggcccgatagtctgaaggtatcagact 810  
 Db 1763 TCAAAAGAGCTCATCTCTGAGCAGTGTGTGGCCCGGTATGCTGAAGGTATCGAGCT 1822  
 Oy 811 gtgccaacttgcagagaagcgttgaacttgaacttggagagataactgtaata 870  
 Db 1823 GTGGCCATTAAGTCAAGAGCGCTGACAAATTAAGCTTTGGAGAGAAATTAAGTGAATA 1882  
 Oy 871 gcccacacatgcatltagaagaacatcgcttggcagccccccagacccagatgtagtccaaagcc 930  
 Db 1883 GCCCCCATATCGTTGGAGAACATCGTTGGCAGCCCCCAGACGAGATGAGACCAACGCGC 1942  
 Oy 931 cgcatagcacacatacaagccgtcttcaacagaagaggttcaacttgcctccacaagcc 990  
 Db 1943 CGCATGACCATATCAAAAGCTGCTTCTACAGAGAAAGGTGACGTTCCCTCCACACGCC 2002  
 Oy 991 gcttcaacacttgcacacttctcgtcgaagaagacttgaatgaacagttgactcactaagttg 1050  
 Db 2003 GCTCTCAACCTGCGACACTTCTGCGCGAAGAGACATGATGAACCACTGACTCATGATGCG 2062  
 Oy 1051 catcaactatgtatgtagagagacttgggttcgcgaagagaccttacaagaataccgctgact 1110

Db 2063 CATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACT 2122  
 QY 1111 ggagaaatgctaaacctggttacctgaaggaagcagctatggtgtaagtaagtaagatg 1170  
 Db 2123 GGAAGAACTGTTAACTGTTGCTGACGAGGAAGAGCTATGTGGTGAAGGTAAGAGAGT 2182  
 QY 1171 gctgggacgagcagtggtgagacggaacccagcagatctggtccagcagcagcgcggaaga 1230  
 Db 2183 GCTGGGGCGGGGTGTGGAGGGAGCCCGACAGATCTGGGCCAGACCTTGGCCGGAAGGA 2242  
 QY 1231 acttcagcgcaaaagcgtgagctcatgacctcaagcagcagcagcagcagcagcagcagc 1290  
 Db 2243 ACTTCAGACACAAAAGGCTGAGCTCATGCCCCCTCACGCACACTTGGGGCTGGCCGGAAGG 2302  
 QY 1291 aaatccataacatttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1350  
 Db 2303 AAATCCATTAACATTATTAACAGACAGCAGGTATGCTTTGCGACTGACACGTAACAGGG 2362  
 QY 1351 gccatcataaacaagggggttgcttactcagcagcagcagcagcagcagcagcagcagcagc 1410  
 Db 2363 GCCATCTATAAGCAAAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAG 2422  
 QY 1411 gaaatcctaagcctatagaagccttaccatttgccaaaagcagcagcagcagcagcagcagc 1470  
 Db 2423 GAAATCTTAAGCCTATTAGAACCTTACATTTGCAAAAAGCGTAGCTATTATACACTGT 2482  
 QY 1471 ccttgagatcagaagaagcagaatctcatatctagagggagcagcagcagcagcagcagcagc 1530  
 Db 2483 CCTTGACATCAGAAAGGCAAGATCCCATTCAGAGAGGAAACAGATGCTGACCGGGTT 2542  
 QY 1531 gccaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1590  
 Db 2543 GCCAAGCAGCGAGCGCCAGGGGTGTTAACTTCTGCTTATGATAGAAACACCCAAAGCCCA 2602  
 QY 1591 gaaacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1650  
 Db 2603 GAACCCGAGCAGACAGTACACCTTAGAAGACTGGCAGAGATTAATAAAGATAGACAGTTCC 2662  
 QY 1651 tcttgagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1710  
 Db 2663 TCTTGAGCTCCGGGAAGGAGCTGCTATACCTCAGATGGAAGAAATCTGCCCCACAA 2722  
 QY 1711 gaaaggttagaatactgccaacagatatacctgcttaacccaacagcagcagcagcagcagc 1770  
 Db 2723 GAAGCGTTAGATATGTCCAAACAGATACATGCTTAAACCACTAGGAATTAACACTG 2782  
 QY 1771 caagcagttgctcagaacatcccttcatatgcttctgaagcagcagcagcagcagcagcagc 1830  
 Db 2783 CAGCAGTTGGTCAGAAACATCTCTTATCATGTTCTGAGGCTACAGAGAGTGCTGATTCG 2842  
 QY 1831 gtgtgcaaacatgtgtgctgctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1890  
 Db 2843 GTGTGTAACACACTGTGTGCTCCCTGCACTGCTGTAATGCTAATCTTCCAAATAACCTCCA 2902  
 QY 1891 ggaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1950  
 Db 2903 GGAAGAGACTTAAGGGAAGCAACCCAGGGCTCAGTGGGAAGTGACTTCACTGAGGTA 2962  
 QY 1951 aagcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2010  
 Db 2963 AAGCGGCTTAATATCGGAACCAATATCTATGTTGTTTGTAGACACCTTTTCAAGGTGG 3022  
 QY 2011 gtaagagcctatactactaagaagaagcagcagcagcagcagcagcagcagcagcagcagc 2070  
 Db 3023 GTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGTGGCTAAGAAATACTGAG 3082  
 QY 2071 gaaatttccaaagatttgtaatacctaagtaataagtcacagcaatgctcagcagcagcagc 2130  
 Db 3083 GAAATTTTTCACAAAGATTGGAATATCTAAGGTAATAGGTCAGAGCATGTGTCAGCTTTC 3142  
 QY 2131 gttgccaggttaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2190

Db 3143 GTTGCCCAAGTAAGTCAGGAGCTGGCCCAAGATATTGGGAGTATTGGAACATTCATTGT 3202  
 QY 2191 gcatacagaccccaaaagcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2250  
 Db 3203 GCATACAGACCCCAAAAGCTCAGSACAGGTAGAGAGATGAATGAACATTAAAGAGAC 3262  
 QY 2251 cttaac 2255  
 Db 3263 CTTAC 3267  
 RESULT 13  
 AX052636  
 LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 21 from Patent WO0071726.  
 ACCESSION AX052636  
 VERSION AX052636.1 GI:12226826  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 porcine endogenous retrovirus.  
 porcine endogenous retrovirus  
 Viruses: Retroviral viruses: Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 6076)  
 REFERENCES  
 Federoff, M.J.  
 Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 Patent: WO 0071726-A 21 30-NOV-2000;  
 Mayo Medical Ventures (US)  
 LOCATION/Qualifiers  
 FEATURES  
 source  
 1..6076  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 1613 a 1512 c 1609 g 1342 t  
 ORIGIN  
 Query Match 24.0%; Score 797; DB 6; Length 6076;  
 Best local similarity 98.7%; Pred. No. 0;  
 Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
 QY 31 tctcgtgagattacacccacatagcagcagcagcagcagcagcagcagcagcagcagcagc 90  
 Db 3577 TCTGCTGATGATTCACCCCACTAGCCACATCTTTGCTTGAATGGAGATCCA 3636  
 QY 91 ggtacggaagaacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 150  
 Db 3637 GGTACGGAGAAACCGGAGCTCAGCTGAGACCCAGCTGCCCCAAGGTTCAAGAACTCC 3696  
 QY 151 ccgacacatcttgacgaagccctacacagcagcagcagcagcagcagcagcagcagcagcagc 210  
 Db 3697 CCGACCATCTTGTAGCAAGCCCTACACAGAGACTGGCCAACTTCAGATTCACACCT 3756  
 QY 211 caggtgacccctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 270  
 Db 3757 CAGGTACCCCTCTCAGTACGCTGATGACCTGCTTGGCGGAGCCCAACCAAGAGAC 3816  
 QY 271 tcttgaaagctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 330  
 Db 3817 TCGTTAAGAGGACGAGAGGACTACTGCTGGAATTTGCTGACCTTGAAGGCTACAGAGCTCT 3876  
 QY 331 gctaaagagccagatcttgacagagagagcagcagcagcagcagcagcagcagcagcagcagc 390  
 Db 3877 GCTAAGCAAGCCCAAGATTGCAAGAGAGAGATTAATCTTGGGAGGATCTTGGCGGAG 3936  
 QY 391 gggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 450  
 Db 3937 GGGCAAGCAGTGTGCTGACGAGGACAGGAGAAACTGTAGTCAATACCTGCGCCCAAC 3996  
 QY 451 acagccaaacagctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 510  
 Db 3997 ACAGCCAAACAGTGTAGAGAGTTTGGGAGACAGCTGCAATTTTGCAGACTGTGGATCCCG 4056  
 QY 511 ggttctgagcctlagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 570

[illegible]

Db	5137	GAACCCGAGACAGTACACCTTAGAAGACTGGCAAGAGATATAAAAAGATAGACCAATTC	5196
Qy	1651	tcctagaatcccggaaggagacctctctataactatataaggaaagaaatctgccccacaa	1710
Db	5197	TCCTAGACTCCGGAAGGAGCTCTATACCTCAGATGGGAAGAAATCCTGGCCACAA	5256
Qy	1711	gaagggttagaatatgctccaaagatatacgtctctaaaccacttagaactaaacactg	1770
Db	5257	GAAGGTTAGAAATATGTCCAAAGATACATCGTCAACCCACTTAGAACTAAACCTGG	5316
Qy	1771	caagcagttgctcagaacatccccttaacatgctctcgttaggtctacagagtggtctacg	1830
Db	5317	CACGAGTTGGTCAGAACATCTCCTTATCATGTCTAGAGCTACAGAGAGTGGCTGATTCG	5376
Qy	1811	gtgtctaaacatgctgtgcacctcgcacgtggttaatgctcaatccctcagaatactcca	1890
Db	5377	GTGGTCAAAACATGTGTGCCCTCCAGCTGGTTAATGCTAACTTCCAGATACCTTCA	5436
Qy	1891	ggaagagactaaggggaagccacccacagcgctcacctcgtggaagttagactcaatga	1950
Db	5437	GGAAGAGACTAAGGGGAACCCACGAGCGCTCACTGGGAAGTGGACTTCACGTAGSTA	5496
Qy	1951	aagccggtctaatacagynaacaatatctatgtgttttctagacacctttcagaatg	2010
Db	5497	AAGCGGCTAAATACGGAACAATATCTATTGGTTTTGTAGACACCTTTCCAGATGG	5556
Qy	2011	gttagagcttatccctactaagaaagagacttaaccgctggtgtgtcctaaagaaatctgag	2070
Db	5557	GTAAGGCTTATCCTCTACTAAGAAAGAGACTTACCGTGCTGTGCTMAAATAACTGAG	5616
Qy	2071	gaaatcttccaaagattgtgaatacctaagatgaataggtctacagaactggtccagcttc	2130
Db	5617	GAATTTTTCAGATTTTGGATTAACCTTAAGGTAATAGGTCAGACAAATGGTCCAGCTTC	5676
Qy	2131	gtgtcccaagtaagtcgaagagactgtgccaagatatgtgggatatgtatgtgnaaactgcatgt	2190
Db	5677	GTGCCAGTAGTGCAGGAGACTGCCAAGATATTGGGATTGATTGAAACTGGATGTGT	5736
Qy	2191	gcatacagaccccaagctctcagacagtgtagagagatgaaatagaacatlaagagacc	2250
D*	5737	GCAATCAGACCCCAAGCTCAGACAGGTAGAGAGAGATGAATAGAACATTAAAGAGAC	5796
Qy	2251	cttac 2255	
Db	5797	CTTAC 5801	
RESULT 14			
LOCUS	AX052634		
DEFINITION	Sequence 19 from Patent WO0071726.	PAT	12-JAN-2001
ACCESSION	AX052634		
VERSION	AX052634.1	GI:12226824	
KEYWORDS	porcine endogenous retrovirus. porcine endogenous retrovirus viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses; I-Mammalian type C virus group.		
ORGANISM	Fedorov M.J.		
REFERENCE	1 (bases 1 to 7362)		
AUTHORS	Methods to inhibit infectious agent transmission during xenotransplantation		
TITLE	Patent: WO 0071726-A 19 30-NOV-2000;		
JOURNAL	Mayo Medical Ventures (US)		
FEATURES	Location/Qualifiers		
source	1..7362		
BASE COUNT	/organism="porcine endogenous retrovirus" /db_xref="taxon:61673"		
ORIGIN	1997 a 1821 c 1881 g 1663 t		
Query Match 24.0%; Score 797; DB 6; Length 7362;			

Best Local Similarity 98.7%; Pred. No. 0;  
Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

[illegible]

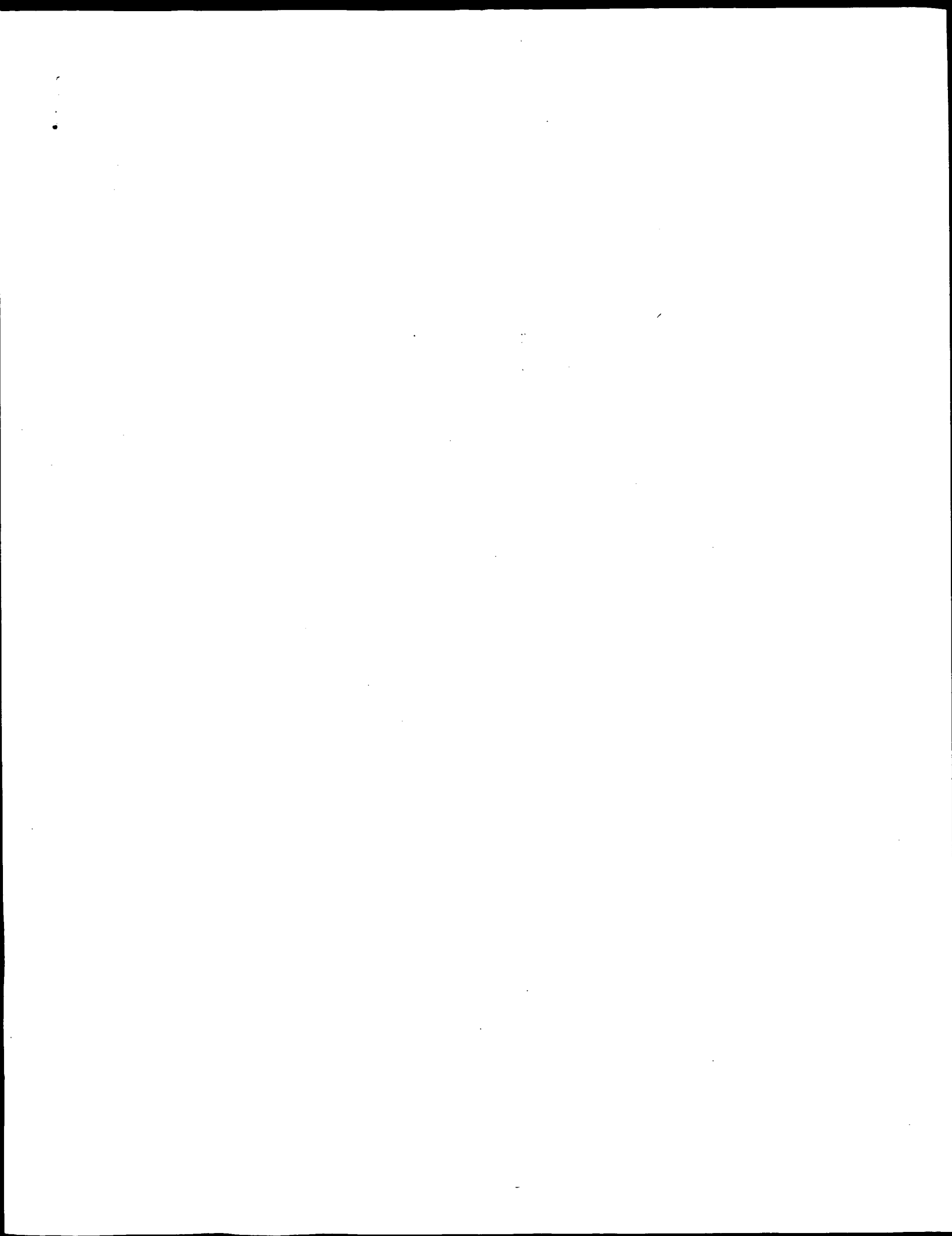
Dh	4662	CATCAACTATTGATTTGAGAGAACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACT	4722
Qy	1111	ggagaagatgctaacctcgtgttctactgacggaacgaactatgtgttggaagtlaagagatgt	1170
Dh	4722	GGAGAGAGTGTAACTCGTGTCTACTGACGGAAGCAGCTATGTGTGGAGAGTAAGAGGATG	4782
Qy	1171	gcttgaggcgacgagtgtgtgaacgggaacccgcgaactctgggcacgaacgctcccgaaaga	1230
Dh	4782	GCTGGGGGGGGGTGTGTGACGGGAGCCCGACGATCTGGGCGCAGCGCTGGCGGAGGA	4842
Qy	1231	acttcaagcgcaaaagcctgagctatgtatgtgacctcaacgaaccttgcgtgttcgcaaggg	1290
Dh	4842	ACTTCAGCACAAGAGGCTGACCTCATGGCCCTCAGCGAACCTTTGGGCTGGCCCAAGG	4900
Qy	1291	aaatccataaacttllatacggacgacgatagtacctttgcagctgcacagtlacacggg	1350
Dh	4902	AAATCCATAAACATTATATACAGACAGAGGTATGGCTTTTCCGCTCGACACGCTACACGGG	4960
Qy	1351	gccactctaacaagaagggtgtcttaccctcagcagggagggaaataaagaacaaagg	1410
Dh	4962	GCCATCTATAAGCAAGGGGGTTCCTTACTCAGCAGGAGGGAAATTAAGACAAAGAG	5020
Qy	1411	gaaatctaaagcctattagaaacgcttaacatttgcacaaaagagctagctallatacactgt	1470
Dh	5022	GAATTTCTAACCTTATGAAGCCTTACATTTGCCAAAAGGCTACTCTATATACACTGT	5080
Qy	1471	octtggaactgaagaagccaaagctctatcttcttaaggagggaacagatgtgtcaccgggtt	1530
Dh	5082	CCTGACATCTAGAAAGCCAAAGATCCCATATCCAGAGGGACCAAGATGGCTCACCGGGTT	5140
Qy	1531	gccaaagcagcgaagcccgagctgttlaacctctgcctataatagaagaagcccaagcccca	1590
Dh	5142	GCCAAAGCGAGGCCCGAGGGGTAAACCTTCTGCTATGATAGAAACCCCAAGCCCCA	5200
Qy	1591	gaaccgcagcgaacgctaccccttagaagactgtgcagaagatataaaagatagaacagtctc	1650
Dh	5202	GAACCCGGAGCAGCATGTCACCAAGATACATGTGCTTAACCCACTAGGAATTAACACTTG	5260
Qy	1651	ctgtgaactccggaggggagactgtctataactctatctgggaaggaatactctgccacaaa	1710
Dh	5262	TCTGAGACTCCGGAAAGGACCTCGTATACCTCAGATGGGAGAGGAATCCTGGCCCAACA	5320
Qy	1711	gaaggggttagaatacttcccaacagatacatcgtctcaaccacactggaactaaacactgt	1770
Dh	5322	GAAGGTTTAGAATATGTGCCAAGATACATCATGTCTAACCCTAGGAATTAACACTTG	5380
Qy	1771	cagcagttgtgtcacaacatccctctatcatgttctbaagctacacggagatgtgtacacgtg	1830
Dh	5382	CAGCAGTTTGGTCAGAACATCTCCTTATCATGTTCTCAGGCTACCCGAGATGGCTGATTCGG	5440
Qy	1831	gtgtgtcaaacatltgtgtcccttcgcacgtgtgtlaatcgtcaatcctctcagaatatcactca	1890
Dh	5442	GTGTCTAAACACTGTGTGCCCTCGCAGCGTGGTTAATGCTAATCTTCCAGATATCTCCA	5500
Qy	1891	ggaagaagcactaaggagggaacccaacccagcgctcaactctggyaagytagtactcaagagta	1950
Dh	5502	GGAAAGAGACTAAAGGGGAACCCACCAAGCGGCTCATCTGGGAAGTGGACTTCATCGAGTA	5560
Qy	1951	aagccgcgtctaatacaggaacaatactcatatgtgtctttgtlagaaccttllcagaatcgg	2010
Dh	5562	AAGCGGCTAAATATACGGAACCAATATCTATTGGTTTTGTAGACACCTTTTCAAGATGG	5620
Qy	2011	gttagaagcctatccctactaagaagaagactcaaccgtgtgtgtcctaagaanaatactagag	2070
Dh	5622	GTAAGGCTTATCTCTACTAAAGAAAGAACTTCAACCGTGTGGCTTAATAAAAAATCTCGAG	5680
Qy	2071	gaaatcttccaagatttgaataactaagtgatataaggtltagaacaatggtgtccagctcttc	2130
Dh	5682	GAATTTTTCACAGATTTTGGAAATACCTAAGGTAATAGGGTCAAGCAATGTGTCCAGCTTTC	5740
Qy	2131	gttgcccaagtlaagtcaggagctgtgcacaagatatltgggagatgtgattgtgaacatgcatgtc	2190
Dh	5742	GTTGCCAGGTAGTACGAGACTGGCCAGGATATTGGGATGTGATGGAACTGCATTTGT	5800

QY 2191 gcatacagaccacaaagctcagagacagtagagagatagaaacattaaagagacc 2250  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 5802 GCATACAGACCCCAAACTCAGAGAGAGTAGAGAGATGATGAACATTAAAGAGACC 5861  
 QY 2251 cttaac 2255  
 |||||  
 Db 5862 CTATAC 5866  
 |||||  
 RESULT 15  
 AX002802 2462 bp DNA PAT 07-SEP-2000  
 LOCUS AX002802  
 DEFINITION Sequence 1 from Patent WO9853104.  
 ACCESSION AX002802  
 VERSION AX002802.1 GI:9885131  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM Porcine endogenous retrovirus.  
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 2462)  
 Stoye, J.P. and Weiss, R.A.  
 Detection of retroviral subtypes based upon envelope specific  
 sequences  
 REFERENCE Patent: WO 9853104-A 1 26-NOV-1998;  
 MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB); WEISS ROBIN  
 ANTHONY (GB)  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source 1..2462  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 715 a 573 c 575 g 599 t  
 ORIGIN  
 Query Match 23.7%; Score 786; DB 6; Length 2462;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 836; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2459 tcgagttggtgagagcagcagcgtggaagagcgtcggaagagcgtactcagagagagact 2518  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 TCGAGTGGGTGAGGCGAGCGAGCGTGAAGAGCGCTCGGAGGCTTACTCAGAGAGAGACT 60  
 QY 2519 tgcagttccacatcgcttccaaagtltggaatcagltcaltgttagagcgcacgltgacg 2578  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 TGCAAGTTCCACATCGCTTCCAAAGTTGAGATTGATGTAGAGCGCACCGCTGACG 120  
 QY 2579 gaaacctcgagactcgtgtggaagagacttactctgacttttgaccacacaaagcgtg 2638  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 GAAACCTCGAGACTCGGTGAGAGGAGACTTATCTCGTACTTTTGACCAACCAAGGCGTG 180  
 QY 2639 tgaagtcgaagagatcccccacgtgacatgcatcccaagtttaagcgcgcgcacgtc 2698  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 TGAAGTGAAGAGATCCCACTGAGATCCATGCCATGCCAGTTAAAGCGGCGCACCTC 240  
 QY 2699 ccgattcgggtggaagccgaaagactgaanaalcccttaagcttcgcctccatcgcg 2758  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 CCGATTGCGGCTGGAAGAGCCGAAAGACTGAANAATCCCTTAAGCTTCGCTCCATCGCG 300  
 QY 2759 tggttccttactctgtcaataactccctcaagttaagttaagcgttggagagccg 2818  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 TGGTTCTTACTCTGTCAATATCTCTCAAGTTATGTTAAACGCTTGTGAGACAGCCG 360  
 QY 2819 aactccataacccattactctacacctgttacttactgactcgcgttacaggtatlaac 2878  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 AACTCCCATTAACCCCTTATCTCTCACTCGTTACTTACTGACTCGGTTACAGGTATTAT 420  
 QY 2879 attaacagactcaagggagagcctcccttgggagcctgltggcctgaattataltgtcgc 2938  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 ATTAACAGACTCAAGGGAGGCTCCCTTGGGAGACTGTGGCTGAATTATATGTCTGC 480  
 QY 2939 ctgcgataagtaacccctggtcctaagtagcagagccacaccccccgaatgtactcgtgct 2998

Db 481 CTTCGATCAGTATATCCCTGGTCTCAATGACGAGCCACACCCCGATGTACTCGTCT 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 2999 tacgggtttacgtlttgcccaagagaccccaaaatgaagaatltgtggaatcctcag 3058  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 TACGGGTTTACGTTTGGCCAGAGACCCCAATATGAAGATATGTGGAAATCTCAG 600  
 QY 3059 gatttccttgaagcaaatgagcgtgcataactctaaatgaggaaltggaatgagca 3118  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 GATTCTTTTGAAGCAATGAGAGCTGCAATTAATCTTAATGATGGAAATGGCCCA 660  
 QY 3119 gctctcagcaagagagtagtactcttltgttaacaatccaccagttataatcaa 3178  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 GTCTTCACAGCAAGACAGAGTAGTACTCTTTTGTAACTTACCAATCTTACCAATGATATCA 720  
 QY 3179 tttaattatgagcagagagatgagaaagattgcaacagcgggtacaaagatgacga 3238  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 TTTAATTAATGAGCATGAGAGATGGAAGATTGGCAACAGCGGGTTCAAAAGATGTACGA 780  
 QY 3239 aataagcaaatagctgltcatctglttagaacttagacttaaaataagttcaact 3295  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 AATAAGCAATTAAGCTGTATCTGTTAGACCTAGATTAATTAATAAATTAAGTTTCACT 837

Search completed: February 24, 2002, 07:20:59  
 Job time: 27774 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      February 23, 2002, 21:01:29      Search time 10411.9 Seconds
              (without alignments)
              5260.413 Million cell updates/sec
```

```

Title: US-09-171-553B-1
Perfect score: 3320
Sequence: 1 gaattcgcgccgctcgac.....aaaaaaaaaaaaaaaaaa 3320

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: genEmb1.\*  
2: gb\_hgt.\*  
3: gb\_hgt.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_om.\*  
20: em\_or.\*  
21: em\_ov.\*  
22: em\_pat.\*  
23: em\_ph.\*  
24: em\_pl.\*  
25: em\_ro.\*  
26: em\_sts.\*  
27: em\_sy.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_hgt\_hum.\*  
31: em\_hgt\_inv.\*  
32: em\_hgt\_ro.\*  
33: em\_hgt\_hum.\*  
34: em\_hgt\_inv.\*  
35: em\_hgt\_ro.\*  
36: em\_hgt\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3320	100.0	3320	6	A66551	A66551 Sequence 1
2	3252.4	98.0	8849	14	PEN133817	ALI33817 Porcine e
3	3227.2	97.2	8918	14	PEN233656	AL293656 Porcine e
4	3179	95.8	4402	6	AX052635	AX052635 Sequence
5	3175.8	95.7	7362	6	AX052634	AX052634 Sequence
6	2978	89.7	7873	6	AX052638	AX052638 Sequence
7	2916.8	87.9	8196	6	A66552	A66552 Sequence 2
8	2916	87.8	8209	6	A66553	A66553 Sequence 3
9	2915.2	87.8	8132	6	ARI330475	ARI330475 Sequence
10	2912	87.7	8132	4	AF038600	AF038600 Sus scrof
11	2876.4	86.6	8132	4	AF038599	AF038599 Sus scrof
12	2867.8	86.4	8918	14	PEN133816	ALI33816 Porcine e
13	2864.6	86.3	8750	14	PEN133818	ALI33818 Porcine e
14	2864.6	86.3	8763	14	PEN293657	AL293657 Porcine e
15	2857.4	86.1	7808	14	PERY17013	Y17013 Porcine eno
16	2721.4	82.0	4918	6	AX052637	AX052637 Sequence
17	2420.6	72.9	8066	6	ARI30473	ARI30473 Sequence
18	2415.6	72.8	6076	6	AX052636	AX052636 Sequence
19	2414	72.7	6076	6	AX052647	AX052647 Sequence
20	1771.8	53.4	7333	4	AF038601	AF038601 Sus scrof
21	1771.8	53.4	7333	6	ARI330474	ARI330474 Sequence
22	1175.8	35.4	8655	6	ARI336169	ARI336169 Sequence
23	1175.8	35.4	8655	14	AF053745	AF053745 Mus dunni
24	1168.2	35.2	154256	2	AC0733295	AC0733295 Mus muscu
25	1168.2	35.2	170658	2	AL566444	AL566444 Mus muscu
26	1147	34.5	274656	2	AC079538	AC079538 Mus muscu
27	1143.8	34.5	185548	10	AC005743	AC005743 Mus muscu
28	1139.4	34.3	202333	2	AC080144	AC080144 Mus muscu
29	1139.2	34.3	218938	2	AL592224	AL592224 Mus muscu
30	1135.8	34.2	196755	2	AC093350	AC093350 Mus muscu
31	1135	34.2	217034	10	AC0812147	AC0812147 Mus muscu
32	1132.6	34.1	65286	2	AC087800	AC087800 Mus muscu
33	1131.8	34.1	199883	2	AC073736	AC073736 Mus muscu
34	1131	34.1	3482	6	AX002804	AX002804 Sequence
35	1131	34.1	3482	14	PEREN2	Y12239 Porcine eno
36	1130.8	34.1	139317	2	AC074337	AC074337 Mus muscu
37	1128.4	34.0	257383	2	AC079494	AC079494 Mus muscu
38	1127.2	34.0	204394	10	AC079644	AC079644 Mus muscu
39	1126.4	33.9	144805	2	AC079555	AC079555 Mus muscu
40	1126.4	33.9	153567	2	AC087145	AC087145 Mus muscu
41	1126.2	33.9	188406	2	AC079560	AC079560 Mus muscu
42	1125.8	33.9	219294	2	AC015797	AC015797 Mus muscu
43	1124	33.6	181572	2	AC020959	AC020959 Mus muscu
44	1117	33.6	207677	2	AL591952	AL591952 Mus muscu
45	1115.6	33.6	158455	2	AC079471	AC079471 Mus muscu



[illegible]

Qy	3181	taatttggccattggagattggaagaattggcaacagcggtacaaaagatgtacga	3240
Db	3181	TAATTATGGCCATGGAGATGGAAGATTGGCAACGCCGGTTCACAAAAGATGTACGAA	3240
Qy	3241	taagcaataaagctgtcattcgttaagaccataattacttaaaataagttcactaaaa	3300
Db	3241	TAAGCAATAAAGCTGTCATTGCTTAGACCTAGATTACTTAAAAATAAGTTTCACTAAAA	3300
Qy	3301	aaaaaaaaaaaaaaaaaaaaa	3320
Db	3301	AAAAAAAAAAAAAAAAAAAAA	3320
RESULT 2			
LOCUS	PEN133817	8849 bp	DNA
DEFINITION	Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class A, clone 42).		
ACCESSION	AJ133817		
VERSION	AJ133817.1	GI:6688947	
KEYWORDS	env gene; env protein; gag gene; gag protein; pol gene; pol protein.		
SOURCE	porcine endogenous retrovirus.		
ORGANISM	porcine endogenous retrovirus		
REFERENCE	Viruses: Retroid viruses: Retroviridae; Mammalian type C retroviruses: 1-Mammalian type C virus group.		
AUTHORS	1 (bases 1 to 8849)		
TITLE	Czaudena, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and Toenjes, R.R.		
JOURNAL	Molecular Characterization of Human-tropic and Replication-competent Porcine Endogenous Retroviruses		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 8849)		
JOURNAL	Toenjes, R.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY		
FEATURES			
source	location/Qualifiers		
	1..8849		
	/organism="porcine endogenous retrovirus"		
	/proviral		
	/strain="type C"		
	/db_xref="taxon:61673"		
	/clone="42"		
	/note="grown in PK15 (ATCC No. CCL-33) cell line"		
	1..668		
LTR			
misc_feature	506		
	/note="cap site"		
primer_bind	671..688		
	/note="gly4"		
gene	1115..2686		
	/gene="gag"		
	1115..2686		
	/gene="gag"		
mat_peptide	/function="retroviral capsid"		
	/product="gag protein"		
	join(1115..2686,2690..6271,6150..8132)		
	/note="viral polypeptide"		
	/codon_start=1		
	/protein_id="CAB65340.1"		
	/db_xref="GI:6688948"		
CDS	/translation="MGQTVTPPLSLTLDHWTEVRSRAHNTSVQVKKGPMPFCASEWME		
	TNDGMSGEFTNSELTIAVKAIIIFQSGPHPODEPIILTWDLAEPPWVAVPLN		
	KPKRPGRIALGKKNKSHAEKVEPSRIPEIETEPWEPQVPPPIPAQAVKRV		
	PSAPGAIVVEGPAAGTSGRGATPEERDETAIIPLRTYGPMPGGOLQIPQWPFSS		
	ADILNMKTNHPPESEDPORLGLVESLMSHQPTWDDCOQLQITFTTEERILLBA		
	RKNPGADVGRPTOLNEDIMGFLPLRPGMDVTAEGRESLKYQALVAGLRGSRRE		
	TNLAKVEVMWGQNPPEPSVLEFLMEAFRRTPDPNTEQAKASVALAFTGSLDIR		
	KTIQRLGELQAEIRDLVREAEKYVRRTEPEEKREKEREERERRRDRDRQKNT		
	KIIAAIVEGSSRRERDRPKIRSGPQSGNLGRTPIDKQCAVCEKGIHWANCPK		
	KGNKGVPVLLAEEDKGRKSDPLPEPRVITLKEGQVEPLVDVIGARSHVLLQIGKI		

FEATURES	Location/Qualifiers
source	1. .8849

```
/organism="porcine endogenous retrovirus"  
/proviral
```

```
/strain="type C"  
/db_xref="taxon:61673"
```

```
/clone="42"  
/note="grown in PK15 (ATCC No. CCL-33) cell line"
```

```
misc_feature 506 1.668
```

```

/ncore="cap site"
primer_bind 671. .688

```

gene	/note="gly4"
1115.	.2686

```
mat_peptide      /gene="gag"
1115.    .2686
```

```
/gene="gag"  
/function="retroviral capsid"
```

	product="gag protein"
CDS	join(115, .2686,2690, .6271,6150, .8132)

```

/note="viral polypeptide"
/codon_start=1

```

```
/protein_id="CAB65340.1"
/db_xref="GI:6688948"
```

```

/translation="MGQIVTTPPLSLDHWVEVRKRAHNLSTYVKRKGKPNQTFCSSEM
TPDVGMPSEGTNSFII LAVKALIFGTGPGSHPDQEPYILTWQDLAEADPPWVAPWML

```

PSADPGAPVVEGPPAGCTRSRSGATPRTDEIALLPLRTGPPMPCGQLQLPLQWPFS

ADLIYMKIINHPPESEDPQRLTGLVSELMFSHQPIWDDCQQLLQLLTALTEERERILLE  
RKKNVGAGDGRPTQLONEIDMGPIILPRGMDVMTAEGRESLKIYRQALVAGLRASRR

INLA VREVMOGPNPSPV LERLMEAFRRITIPDTS EADKASVALAF IGQSMADI  
KLÖRLEGLOEALRDVRAEKYYIRRETEEEKEDORK EKEEREERERDRQENL  
KTI AUMDQGGDD

KGNKGPVLALIEDKDGRRGSDPLDEPRVTLKVEGQVPELVDTGAHSHVLLQPLGK

KEKSWMGATGQROYPWTTTTRTVDLGVGRVTHSEFLVECEPVLGRLDLTKKMAQI  
 SPEQREVSNNKPIITVLQDDEIRLYSPQVKKDDIOQMILQFQPMAPATAGM  
 LAKOVPVQIKASATPVSVROYPLSREARBEIIMHVOYLQGLIVVQSPMPTPL  
 LPRKPGTNDYRQVODIREVNRKQVODIHTVNPVNLISALPERMHWLIDLKAEF  
 CLRHPTISQPLFAFEWRDQRTGOLMTWRLPQCKNSPTIFEDALHMDVILNFKIOH  
 POYTLLOVYDDILLAGATKODCELETKALLLESLDGRASKAKKOICREVTYIGTS  
 LRGGQWLEARKTIVQIPAPITTAQVEPELIGTGFCEMLIPGATVLPQALPILKE  
 KGEFSAWPEHQAQFVDAIKKALLSAPALAPDYKPTLVLDKRGVARNVLOTGLPW  
 RRPVALISKKLDPVASGWPCLKAIAVAIIVKADKLKLGONITVIAVHALENTVRO  
 PPRMNTNARMHYOSLLTERVETAPALNPATLPEDETEPYTHOCHOLEIBETG  
 VRKDLIDPLGTGVLWMTDSSSYVEEGRMAGAAVVOGRTTMSAIDEGNSAKAE  
 LMAIJOALALBSKSNITYTDSRYAFATRAHVGAITIKOGILLTSGRETKNEELSL  
 LEALHLPRKALITHCPGHOKAKDLISRGQMDRVAKQAQAVNLPIITETPKAEPR  
 ROYTLEDMOEIKKIDQSETPEGTQSYGKIBLPKBELEVOOIHRLTGHETHLQ  
 OLVRTSPYHVLIPGVADSVKHCVCQVLANPSPRIIPGKRLSGHPAHHEVETE  
 VKPAKGNKYLLVEVDTSFSGWEAYPTKETSTVAKKILEELIPREPKYIGSDNG  
 PAVNAVOSGLAKIIGIDMKLHCAYRPOSSGOVERNRIKTEITLKTETINDMA  
 LLEFVLEFRVANTPGQGLPPELTYLGGPPLVEIASVHASADVLDSOPLSRKALEW  
 KORAMQLEAVSGEGDLOVPHRFQVDESVYRRHRAAGLLETRMGPIVLLTTPJAV  
 KVERIPWIMHSHVKKPAPPDGSMKAEKENDLKLRLHVVVSYNNSSMHPITLSR  
 HLPTRGKPKRLKIPLSFASIAMFELITTPQVNGKRLVDSNSHKLPLSLTLNDG  
 TGINISTGGEAPLGTWMELEYVCLRSVTPGLNDATPDVLRAYGVYCPGPNNIE  
 YCNPQDFECCOMCTTSDNGMKNPVSODRVSYSVNNPSTSYNOFPHGRMKMO  
 QRYOKRVNRKQISCHSLDITYKISTEKGKODENOKVNGISKEITVYGGSKRKS  
 VLTRIRLETOMEPPVAIGPNKGLABQPPIOEQRPSPSDYNTTSSVTEPEPITI  
 KTGAKLESLIOGAFALINSTPEATSSWCLASGPPYEGARKGNKVRKEHNDQ  
 TWGSONKRLTLEVSQKTCIGWPSHOMLCHNTAFNNTSEQVLEVGDDWMACT  
 GLTPCVSTLWOTKDFCVMQOIVPVVYPERKALDKYNNRPREDPILSTLAVM  
 LGIGVAVAGTGTGAALITGPOOLERKSLHRIHVEDLOALEKYSNPEESTISSEV  
 VLONRGDLFLFLKRGGLCVALKRECCFVYDSGAIROSMKRLRLERRRERERADO  
 GMEGFNRSFMHTTLLSALTGPLVLLLTGVLPCILNREFAVERVSAVQIVNLR  
 QOQGLLSQGEHDL"

mat.peptide

/gene="pol"  
 /product="retroviral protease and polymerase"  
 /function="pol protein"

gene

mat.peptide

/gene="pol"  
 /gene="env"  
 /function="retroviral envelope"

gene

LTR

POLYA<sub>1</sub>signal

BASE COUNT 2355 a 2191 c 2281 g 2022 t

ORIGIN

Query Match 98.0%; Score 3252.4; DB 14; Length 8849;  
 Best Local Similarity 99.3%; Pred. No. 0;

Matches 3278; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

QY 19 aacgagtcctcttctgagagattacaccctagccaaccatttttgccttcgaa 78  
 Db 3497 AAGAGATCCTCTTCTGCGGAGATTACACCCACCTAGCCAACTTTTGGCTTCGA 3556  
 QY 79 tggagatcaggtacggaagaaacccgagcagctcagcagcagcagcagcagc 138  
 Db 3557 TGGAGATCCAGGTAGGGAAGAACCCGGCAGCTCAGCCGACCTGCCCAAGG 3616  
 QY 139 ttcagaagactcccgacacatcttgaagaaacccctacacaggaagcctggccaactcag 198  
 Db 3617 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGGAGCCTGGCCAACTTCAAG 3676  
 QY 199 atcaaacacctcaggtgacctccctccagtagctgagctgcttctggggagagcc 258  
 Db 3677 ATCCAACACCCCTAGGAGACCTCTCTCAGTACGTGATGACTGCTTGGGGGAGAGCC 3736  
 QY 259 accaagaagagctctagaaggtacgaagcactactcgtgaatctgtcgaagcagc 318  
 Db 3737 ACCAAGACGAGCTGCTTAGAAGGTACGAAGGACACTGCTGTAATTTGCTGACCTAGGC 3796

QY 319 tacagaacctctgtctaaagaagcccaagatttcagaagagaaggtacaactctggggtac 378  
 Db 3797 TACAGAGCCTCTGTCTAAGAAAGGCCACAGATTTCAGAGAGAGAGTAACTACTGGGGTAC 3856  
 QY 379 agtttcggggcgcgagcagatgcttcgagagcagcaggaagaaactgtgttcagata 438  
 Db 3857 AGTTTCCGGGGCGGGCAGCAGGATGGCTGAGCGGAGCAGGAAGAAAGCTGTGTCAGATA 3916  
 QY 439 ccggccccacacacagcacaagaagtgagagagtttttggggagacgctggatttgcaga 498  
 Db 3917 CCGGCCCAACACAGCCAAACAGTGAAGAGTTTGGGACAGCTGGATTGGCAGA 3976  
 QY 499 ctgtgattccgggggttgcgaccttaacagcccaactcaccgctlaaccaaagaanaaa 558  
 Db 3977 CTGTGGATCCGGGGGTTGGCAGCTTACAGGCCCACTACCCGCTAACCAAGAAANA 4036  
 QY 559 gggggatctctctgggctctctgagcacaagaagcatttgatctataaagaagccctg 618  
 Db 4037 GGGGAATTCCTCTGGGCTCTGAGCAGCAGAGGCAATTTGATCATCAAAAAGGCCCTG 4096  
 QY 619 ctgaagcgaacctgctctgagccctccctgaagtaaaccttaaccttatgtgat 678  
 Db 4097 CTGAGGCACTGCTCTGAGCTCTCCCTGAGCTAACCTTAACCTTTATGTGAT 4156  
 QY 679 gaggctlaaggagtagcccgagagattttaaaccacaacccctaggacacatggaggagact 738  
 Db 4157 GAGCGTAAGGAGGTAGCCCGAGAGATTAAACCCAAACCTTAGGACATGAGAGAGACT 4216  
 QY 739 gttgctactcgttcaagaagcttgatccctgtagcagtggttggccgtatgtctgaag 798  
 Db 4217 GTTGGCTACCTGTCAAAAGAGCTTATGCTGTAGCAATGAGTGTGGCCGTATGTCTGAG 4276  
 QY 799 gctatcgagcgtgtgccaatctggttcaagaagcgttgacaattgactttggagcagat 858  
 Db 4277 GCTATCGCAGCTGTGTGGCCATACTGTCAAGAGAGCTGACAAATTGACTTTGGGACAGAT 4336  
 QY 859 ataactgttaatgcccccatatgcatlttgagagacatcglttcggcagccccagacagatg 918  
 Db 4337 ATAACTGTAAATAGCCCCCATATGATGGAAGAACTGTCGGAGCCCCAGACGATAGG 4396  
 QY 919 atgaccaaagccgagatataccactataaagccgctcttcaagaagaaggttaacttcc 978  
 Db 4397 ATGACCAAGCCCGCATGACCCACTATCAAGACCTGCTTCTACAGAGAGGTCACCTTTC 4456  
 QY 979 gctccacagccgctctcaacccctgcaactcttctgcttgaagaagactgaagacagctg 1038  
 Db 4457 GCTCCACAGCCGCTCTCAACCCCTGCACTGTTCGCTGAAGAGACTGATGAACCAAGT 4516  
 QY 1039 actcatgattgcatacaactatgtatgaagagactgggttcgcaagagacttacaagac 1098  
 Db 4517 ACTCATGATTGCAATCAACTATTGATGAGAGACTGGGGTCCCAAGGACTTACAGAC 4576  
 QY 1099 ataccgctgactggagagagttctaacctggttcaacttgaacgaagaagctatgtgtgaa 1158  
 Db 4577 ATACCGCTGACTGGAGAGTGTACTGCTGTTCACTGACGAGAGAGCTATGATGTGAA 4636  
 QY 1159 ggttaagaagatgctctggggcgagctgtgtgaagcggacccagatcttggccaagcagc 1218  
 Db 4637 GGTAAAGAGATGGCTGGGGCGGAGTGTGAGACGGGACCGCACAGATCTGGGCCAGCAGC 4696  
 QY 1219 ctgcccgaagaacttcaagcgcaaaagcttgagctcatatgacctcaagcaagcttgcgg 1278  
 Db 4697 CTGCCGGAAGAACTTACGCCCAAAAGGCTAGGCTCAAGGCTCAGCAGCAAGCTTTGGG 4756  
 QY 1279 ctggcgaagaagaaatcatataaactataacgaagaagatgacttcttggagctgca 1338  
 Db 4757 CTGGCCGAAGGAATTCATTAACATTATACGGAACACAGATATGCTTTGGCACTGCA 4816  
 QY 1339 cactfacaaggggcatctataaacaagaaggggttcttactctcagcagggaggaata 1398  
 Db 4817 CACGTACACGGGGGCATCTTAACCAAAAGGGGTTGCTTACTCTAGCAGGAGGAATA 4876

QY 1399 aagaacaaagaggaatcttaagccttattagaagccttacttgccaaagagcctagct 1458  
 DB 4877 AAGACAAAGAGGAATTTAAAGCCTATTAGAGCCTTACATTTGGCCAAAAGGCTAGCT 4936  
 QY 1459 attatacaatctcttgacatcagaagaacaaagatcatalactagaggaagaccagatg 1518  
 DB 4937 ATTATACAGTCTCTGGACATCGAAGAACCAAGATCTCATATCTAGAGGAAACGATG 4996  
 QY 1519 gctgacccggttgcaagcagcagccagcgtctgtaacctctgctcttaataagaacg 1578  
 DB 4997 GCTGACCGGGTTCCAGGACGGGACGCGGCTTACCTTGTCTATTAATAGAAAGC 5056  
 QY 1579 cccaagccccaggaacccagagacagctacacccctagaagactgtaagaagataaagaag 1638  
 DB 5057 CCCAAAGCCCCAGAACCCAGACACAGTACCCCTTAGAAGACTGGCAAGAGATAAAAAG 5116  
 QY 1639 atagacagtlctcttgagactccggaagggagcctgctataacctcatalatggaagaatc 1698  
 DB 5117 ATGAGACAGTCTCTGAGACTCCGAGGGGAGCTGCTATACCTCATATGGGAAGAAATC 5176  
 QY 1699 ctgcccacaaagaaggttagaataatgltccaaacataatctgttaaccacccataga 1758  
 DB 5177 CTGCCCCACAAAGAGGTTAGATATGTCCAAACAGATACATGCTCTACCCCTTAGA 5236  
 QY 1759 actaaacactgacagaggttgacagaacatccctatcatgttctgaagctacagga 1818  
 DB 5237 ACTAAACACCTGACAGAGTTGGTACAGACATCCCTTATATGTCTGAGAGCTACCAAGA 5296  
 QY 1819 gtggtctactcgtgtgcaacacatgtgtgcccctgcccagctgtgtaagttaacctcc 1878  
 DB 5297 GTGGCTGACTCGTGGTCAAAACATGTGTGCTCCCTGCGAGCTGGTTAATGCTAATCCCTC 5356  
 QY 1879 agaatctcccaagaaagaactaaagggagagccacccagcgtctcatctgaggaagtgac 1938  
 DB 5357 AGAATACCTCCAGAAAGAACTAAAGGAGGACCCAGCCCTCACTGGAAGTGAGC 5416  
 QY 1939 ttactaggttaaagccggttaatacaggaacaaatcatatgtgttctgtagaac 1998  
 DB 5417 TTCTACTGAGTAAAGCCGCTAAATACGGAACAAATATCTATTGTTTGTAGACACC 5476  
 QY 1999 ttctcagatggttagaggtctatctactaagaagaagactcaacgctgtgtgctaa 2058  
 DB 5477 TTTTCAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTCAACCGTGGTGGCTAA 5536  
 QY 2059 aaatctctgaggaatcttccaaagtctggaatctgaaggaatgaaggttagagaaat 2118  
 DB 5537 AAAATACTGGAGAAATTTTTCGAGATTGGAAATACCTAAGGTAATCGGGTACAGCAAT 5596  
 QY 2119 ggtccagcttctggtcccaagtaagtcagagactggtgccaagatatggtggtatctg 2178  
 DB 5597 GGTCCAGCTTCTGTTGCCAGGTAAAGTCAAGGACTGGCCAAAGATATGGGGATTGATTGG 5656  
 QY 2179 aaactgcatgtgcatcacagcccaaaagctcagagcaggttagagagatgaatagaac 2238  
 DB 5657 AAACGTGATGTGATACAGACCCCAAAAGCTCAGAGACAGTAGAGAGATGAATAGAAC 5716  
 QY 2239 attaaagagaccccttaccacaaatgtgacacagagactggtcattatgtttgagtctc 2298  
 DB 5717 ATTAAGAGACCCCTTACCAAAATTAACCAACAGAGACTGGCATTAATGAATGGATGGCTCTC 5776  
 QY 2299 ctgccccttgctgctttttaggttagaggaacaccccttgacaggttctggcctgaacccct 2358  
 DB 5777 CTGCCCCCTTGTGCTTTTAAAGGTAGAGAACCCCTTGACAGTTTGGCTGACCCCTAT 5836  
 QY 2359 gaattgctctacggggagaccccccgttgacagaatctgctcttgacatagtgctgat 2418  
 DB 5837 GAATGCTCTACGGGGAGACCCCCCGTGTAGAAAATGCTCTGTACATAGTGTGAT 5896  
 QY 2419 gtgctgcttccagccttctgtctaggtcagaaggtcaggtgtgtgtaggcaagca 2478  
 DB 5897 GTGCTGCTTCCAGCCCTTGTCTTCTAGGCTCAAGGGCTCGAGTGGGTAGGCAACGA 5936  
 QY 2479 gcgtggaagcagctccgggagagcctactcagc---aggaagactgcaagttccacatgc 2535

DB 5957 GCGTGAAGCAGCTCCGGGAGGCTACTCAGAGAGAGAGACTGCAAGTTCACATCGC 6016  
 QY 2536 ttccaagttgagatctgctatagttagagccacggtgaggaagaacctgagacg 2595  
 DB 6017 TTCCAAAGTTGAGATTCAGTCTATGTTAGACGCCACCTGCAAGAACCTGGAACCTCG 6076  
 QY 2596 tgaagagaccltatctgtaacttltgacacacacaaagcgtgtgaaggtcgaaagac 2655  
 DB 6077 TGAAGGGGCCCTTATCTGTAAGTTTGAACCAACCAACGCGTGGAAGTGAAGTAATTC 6136  
 QY 2656 cccacttgatccatgcatcccaagttaaagccgagccactcccgatctggggttgaa 2715  
 DB 6137 CCCACTGTGATCCATGATCCACGCTTAAGCGGGGCCACTCCGATTCGGGGTGAA 6196  
 QY 2716 gccgaagaactgaaatccccccttaagcttgcccctcctcagtggttctactctg 2775  
 DB 6197 GCGGAAAGACTGAAAATCCCTTAAAGTTCGCTCCATCCGCTGCTTACTCTGTC 6256  
 QY 2776 aataactcctcaagtltaagtaagcgttgtagacagccgaactcccaaaacct 2835  
 DB 6257 AATTAACCTCTCAAGTTAATGTAAGCGCTTGTGACAGCCGAACTCCATTAACCTT 6316  
 QY 2836 atctctacactggttacttactgactcggtaacaggtatataataacagcactcaag 2895  
 DB 6317 ATCTCAACCTGTTACTTACTGACTCCGTAACAGATTAATTAATTAACAGCATCAAG 6376  
 QY 2896 ggaagctcccttgaggagcctgtgtgccaatataatgltcccttgacatgaatcc 2955  
 DB 6377 GGAAGCTCCCTTGGGGACCTGGTGGCTGAATTAATATGTCTCTCCATCAGTAATCC 6436  
 QY 2956 tgtctcaatgacagccacacaccccgatgtactcgtgtgctacaggttactgct 3015  
 DB 6437 TGGTCAATGACAGGCCACACCCCGATGTACTCCGCTTACGGGTTTACGTTTG 6496  
 QY 3016 cccaagaccccaataatgaagaatattgtgaatctcgaagttcccttgaaaca 3075  
 DB 6497 CCCAGAGCCCCCAAAATGAAGAAATTTGTGAAATCTCGAATTCCTTTTGAAGCA 6556  
 QY 3076 atgagactgcatcaacttctcaatgagatggaatggaagctcctcagcaagacag 3135  
 DB 6557 ATGGAAGCTCATTAATCTTAATGATGGAATGGAATGGCAAGTCTCAGACAGACAG 6616  
 QY 3136 agtaagtlactccttltgtaacaatctcacosgtlataatcaatlaattatgccaag 3195  
 DB 6617 AGTAAGTACTCTTTTGTAAACATCTCCAGTTATTAATTAATTAATTAATTAATTAAT 6676  
 QY 3196 gagaatgaaagattgcaacacggtgtacaaaagaatgagaaatgaagaataagctg 3255  
 DB 6677 GAGATGAAAGATTTGGCAACAGCGGTACAAAAGATGACGAATTAACGAATTAAGCTG 6736  
 QY 3256 tcaatgcttagactgattacttaaaatagttcactaaaaaaaataaataa 3315  
 DB 6737 TCATTCGTTAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6796  
 QY 3316 aa 3317  
 DB 6797 TA 6798

RESULT 3  
 PEN293656 8918 bp DNA VRL 30-MAY-2001  
 LOCUS Porcine endogenous type C retrovirus proviral gag gene, pol gene  
 DEFINITION  
 and env gene, class A, clone 58.  
 ACCESSION  
 AJ293656  
 VERSION  
 AJ293656.1 GI:14275837  
 KEYWORDS  
 env gene; envelope; gag gene; group specific antigen; pol gene;  
 polymerase; polypotein.  
 SOURCE  
 porcine endogenous type C retrovirus.  
 ORGANISM  
 Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.  
 REFERENCE  
 1 (bases 1 to 8918)

FEATURES	Location/Qualifiers
source	1. .8918

CDS

/gene="gag"  
 /codon\_start=1  
 /product="group specific antigen"  
 /protein\_id="CAC39616.1"  
 /db\_xref="GI:14275838"  
 /translation="MGTQVTPPLSTLIDHWTEYKSHANI,SYQVKKGPQWQFCVSEWTFEDGWSDEGFENSELLAAKAIIFQTGSHDDEPYILTMODLAEDPPVAVKPKWLNKPPRGGRIALVLEGKNNKSHAEKATPSPHIYPELEEPPAPPEOSVPPPYAQAQARSPPAPGPAATGAGPAAGTGRSRGATPRTPEITLPLRTGPGPIPGGOLQPLQWPERSADLVNMTNHPPEFEDPORLTLGVESIMESHQPTMDCOOLLQTLPTTEBKRILLERKNGAGDGRPTOLONEIDMGPIITRPGMDYPMABREKATYRQALVAGLNGSRRTNLKVEWQWQNPPEPSVYLERIMAFNRRFLPDPDTLSQAASKALVAITQSLALDIRKKIQLRGLOEAEHLHYKAEKVIYKRTTEERQKREKEEEERNNRQKRLNLTITLAAVEGKSNREREDFRKITSRPGSGNLGNRTPLDKDCAAYCKEKGHWADDCPKKNGKGLVALLAEEDKD"  
 2728..6309  
 /gene="pol"  
 /gene="pol"  
 <2728..6309

gene  
CDS

16185..8149  
6185..8149  
/gene="env"  
/gene="env"  
/codon\_start=1  
/product="envelope"  
/protein\_id="CAC39618.1"  
/db\_xref="GI:14275840.1"  
/translation="MHPTLSRRRLPIRGKPKRLKIPLSFASIAWFLTSLTPVNGF  
RLVSPNSHKPLSLTALTDSTGTINISTGEAPLGTWPELVGLASVPIGNDQ  
TLPVULRAYGFCYVCPGPNNNEYCGINPDGFCQMSCTSDNGWKKPVSODPRYSY  
FVNPDTSNOENFNGSRKMKDQQRVDKRVNRKQJSCSHLDLYLKISTEKGENIT  
KNVNCISMGILYVYGSGRKKKSVLTIRLITGTOMEPPVAIPNKGLEAQGPIDQIR  
SPNDSYNTITGSVPTEPNITIKTGAKLFSLIOGAFALNSTPEAVSSCMWCLASGP  
PYLYGMARGGKGFVNTKREHRCQCTGSSONKLTLTVEGSGCTIGVPSHOLCHKEVE  
FNRTSEQYLPGYDWRMACNTGLTPCVSTLTFTNOTDEFCWMOIVRVRYVYKATV  
DEVYRYRNPKRREPILSLIAWMLGLGAAGCPTAALTITGPOLBGLSLMHLIVL  
DLQLEKSVSLSESLTSLSEVYLQNRGDLTLFKBGLCVALKERCCFLYVDHSGAL  
RDSMSKTLRELRKHKERKEAGQGFEFGMFKNSPWTTLLSLTGLPVLITLTLTGPGC  
LINFEVAFVREQQSAVRIMVLRQOYGLPS"

LTR  
8216..8918  
8810..8815

polya\_signal  
BASE COUNT 2435 a 2253 g 2056 t

ORIGIN

Query Match	97.2%	Score 3227.2	DB 14	Length 8918
Best Local Similarity	98.7%	Pred. No. 0		
Matches 3253	Conservative	0	Mismatches 43	Indels 0
			Gaps 0	
QY 22	gatyccctctctctgcctcgtgagattacaccocactagccaaccacttttgcctcgaatg	81		
Db 3538	GATGCTCTTCTTCTCCGAGACTACATCCCTAGCCACACACTTTTGCTCTGAAATGG	3597		
QY 82	agaatccaggtacgggaagaacccgggcagctaccgcggaccgactgccccaaaggctc	141		
Db 3598	AGAAATCCAGGTGGGGAGACACCGGCGAGCTCACTTGATGATCAGTCGCCCAAGGGTTCC	3657		
QY 142	aagaactccccgaacalcctttgaagaagccctacacagggagccctgycgaactcagatc	201		
Db 3658	AAAAACTCCCGAGACATCTTTGAGGAAGGCCCTACACAGAGACCTGGGCCAACTTAGCATC	3717		
QY 202	caaacacctagtgtagaacctctccagtacgttgatgtagctgtcttgcgcggagccacc	261		
Db 3718	CAACACCCCGAGGTGACCCCTCTCCAGTACGTGGATGATCCCTGCTTGGCGGGAGCCACC	3777		
QY 262	aaacagagactgcttaagaagtlacgaagacactactgttgtaatgtgtcagctagagctac	321		
Db 3778	AAACAGGACTGCTTTAGAGGTACGAAGGCACTACTGCTGGAATGTCTGACCTAGGCTAC	3837		
QY 322	agaagcctctgctlaagaagcgccagatttgcagaagaaggtlaacatactgggggtacgt	381		
Db 3838	AGAGCCTCCGCTTAAGAAAGGCCAGATTGTGAGAGAGAGAGGTAACTATCTTTGGGGTACAGT	3897		
QY 382	ttgcgggggcgggcagcagatgctgtacgggagccggaagaanaaaactgtagtcagatalccg	441		
Db 3898	TTGCGGGGGCGGGCAGGATGGCTGACGGAGGACGCGAAGAGACTGTAGTCAGATACCG	3957		
QY 442	gcccccaaccaagccaaanaaafgagaagggtttggggagcagctgatttgcagactg	501		
Db 3958	GCCCCAACCAACACCAAAACAGTGAAGAGATTTTGGGACACACTGTGATTTTGCAACATG	4017		
QY 502	tggatcccggggttttgcagaccttagcagaccacactacccgtlaaccgaagaanaaagg	561		
Db 4018	TGGATCCCGGGGTTTGGCACTTTAGCAGCCCACTATTCCCACTTAACCAAGAAAAAGGG	4077		
QY 562	ggattctcctgggctctgtagcaccagaagaagcatctgtagtctatacaaaaaggccctgctg	621		
Db 4078	GAAATCTCTCTGGGCTCTTAGCACCAGGAAGGCAATTTGATGATCATCAAAAAGGCCCTGCTG	4137		
QY 622	agcgacactgctctggccctccctcctgacgtaactaaaccttaccccttatgttgatgag	681		

D 4138 AGGCACCTGCTGCGCCCTCCCTGAGTGAATTAACCTTTTACCTTTATGTGATGAG 4197  
Q 682 cgttaaggagtagcccgaggagttttaaccacaccccttaagacatlgagagagactgtt 741  
D 4198 CGTAAGGAGTAGCCCGGGGCTTTTAACCAAACTTAGACCATGAGAGACCGCTGT 4257  
Q 742 gccactctgtcaagaagacttgatccctgttagccaagtgtgttgcccgtagtctgaagct 801  
D 4258 GCGTAACCTGTCAAGAGAGCTGATCTGTAGCCAGTGTGGCCCATATCCCTGAAGGCT 4317  
Q 802 atcgacactgtgacatactgtgtcaagaagcctgacacaaattgaacttttggagagaatata 861  
D 4318 ATGCCACTGTGGCCATACCTGCTGACAGGACGCTGACAAATTTGACTTTGGACAGAAATA 4377  
Q 862 actgtaataagcccccaatgcatltgagagacatcgttcgcagaccccccaagacatgagt 921  
D 4378 ACTGTATATAGCCCCCATGGGTTGGAGACATGTTGGGAGGCCCCAGACGATGATG 4437  
Q 922 accaagcccgcatgacccaatcaaaagccttcctcaagagagagtgacttcgt 981  
D 4438 ACCAAGCCCGCATGAGCCCACTATCAAAAGCCTCTTCTCAGAGAGAGGCTCAGTTCGCT 4497  
Q 982 ccacacacgctctcaaccctgcacactctctgcctgaagaagactatgaaccagtgact 1041  
D 4498 CCACACGCCCTCTCAACCTGCGACCTCTGCTGTAAGAGACTGATGAACCAAGTACT 4557  
Q 1042 catgattgcatcaactatgtatgttagagagactgggtccgcgaagacacttaacacata 1101  
D 4558 CATGATTGCCATCACTATTGATTGAGAGACATGGGGTCCGCAAGACCTTACAGACATA 4617  
Q 1102 ccgctgactgagagaagtgttaacctgttcaactgacggaagacgtatgtgtgtgaaggt 1161  
D 4618 CCGGTGACTGAGAAATGTTAACCTGTTCACTGACGAGGAAGCAGCTATATGTGTGAAGGT 4677  
Q 1162 aagagaatgcttgaggcggaagtgtgagagagacccgaagacatcgtgggcccagcacgtg 1221  
D 4678 AAGAGATGCTGCTGGGCGGGGTGTGTGAGCGGACCGGACGATCTGGGCGCACCGCTG 4737  
Q 1222 ccggaaggaactcaagcgcgaagaagctgagctcatgtgacctcaagcaactgttgcgctg 1281  
D 4738 CCGGAAGGAACCTTACAGACAAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGGCGCTG 4797  
Q 1282 gccgaaggaatccatcaataacatttaacgagacagagtagtgccttgcgactgcacac 1341  
D 4798 GCGGAAGGAAATCCATAAATTTATACAGACAGCAGGATGCTTTGCGACTGCACAC 4857  
Q 1342 gtacacggggccatctataaacaagaagggtgtgcttaactcgaagcgagggggaataaag 1401  
D 4858 GTACACGGGCGCATCTATATAAGCAAGGGGTTGCTTACCTCAGCAGGAGGAAATAAAG 4917  
Q 1402 aacaaagaggaatcttaagacctatagaagccttaactttgccaagaagcctagtatt 1461  
D 4918 AACAAAGAGGAATTTTAAGCCCTATTAGAAGCCTTACATTGGCAAAAAGGCTAGCTATT 4977  
Q 1462 atacaactgttcctgacatcaagaagccaagatctcaatctagaggaacagatgtgct 1521  
D 4978 ATACACTGTCTCTGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAAACAGATGCT 5037  
Q 1522 gaccgggttgccaagcagcagcgagctgttaacctctgctaataaagaagcc 1581  
D 5038 GACCGGGTTGCCAAGCAGGACGCCAGGGGTGTTAACCTTCTGCTATGTATTAACACCC 5097  
Q 1582 aaagcccccagaaccagacagtagtacacctagaagactgtgcaagagataaaagaata 1641  
D 5098 AAAGCCCAAGAACCCGGAGACAGTACACCTTAGAAGACTGGCAGAAATTAATAAAGATA 5157  
Q 1642 gaccagtctctgagactcccgaggagactgctataactcctatagaggaagaatcctg 1701  
D 5158 GACCAAGTTCTCTGAACCTCGGAGGGGACCTGCTATACCTCAGATGGGAGGAATTCCTG 5217  
Q 1702 ccccaaaaagaggttagaatatgtccaacagatatacgtcttaaccacctaagaact 1761  
D 5218 CCCCAAAAGGAAGGTTAGAAATATGTCCACAGATATCATGCTTAACCCACTAGGAACCT 5277

Q 1762 aaacacctgcaagactgtgtcagaacatccccttaactgttcttgaggctaaccaagtg 1821  
D 5278 AAACACCTCAGACAGTGTGTCAGAACATCCCTTATCATGTCTGAGGCTACAGAGATG 5337  
Q 1822 gctgactcgtgtgtcaaacatgtgtgtccctgcagcagctgtgttaatgttaactccaga 1881  
D 5338 GCTGACTCGGTGCTCAAAATGTGTGTGCTCCAGCTGTGTTAATGTAATCTTCCAGA 5397  
Q 1882 ataacctccagaaaagagactaaagggaagcccccagcgctcactgtggaatgnaactc 1941  
D 5398 ATACCTCCAGAAAGAGACTTAAGGGAGACCCACAGGGGCTCCTGGAAGTGAATTC 5457  
Q 1942 actgaggtlaaagcccgctaaataacgaaacaatactatlttgatttltglaacacct 2001  
D 5458 ACTGAGGTAAACCCGGCTTAATACGGAACAATATCTATGTGTTTGTGTACACACTT 5517  
Q 2002 tcaagatgggttagagcttaactcctaagaagaagactcaaccgtgtgttaagaa 2061  
D 5518 TCAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTTCAACCGTGTGCTTAAGAA 5577  
Q 2062 atactggaagaaattttccaagatttgaatataagtaatggttcaagacatgt 2121  
D 5578 ATACTGGAAGAAATTTTCCAAAGATTGGAATTAACCTTAAGGTATAGGCTCAGACATGCT 5637  
Q 2122 ccagcttctgtcccaagltlaagltcaaggaactgccaagataltgaggatgattgaa 2181  
D 5638 CGAGCTTGTTGTTGCCAGGTAGTCAAGGACTGCGCAAGATATTTGGGATTTGATGAAA 5697  
Q 2182 ctgcatltgtcatcagaagcccgaagctcaagacagtagagaga tgaatagaacatc 2241  
D 5698 CTGCAATTGTGCTATACAGACCCCAAGCTCAGACAGTAGAGAGATGAATGAACCAT 5757  
Q 2242 aaagagaccttaccacaaattgaaccaagagcttgatcatatgtatgtatgtctcctg 2301  
D 5758 AAGAGACCTTATCAAAATTTGACCAAGATGAGATGCAATTAATGATTGATGGCTCTCTG 5817  
Q 2302 ccccttgtgcttlttaggttgagagacaccccttgacacgtlttggtctgaaccctatgaa 2361  
D 5818 CCTTTGTGCTTTTATAGGGTGAAGAACCCCTGAGACGTTTGGCTACCCCTCATGAA 5877  
Q 2362 ttgctctacgggggaaccccccgcttgcaagaatltgcttgcataagtgtcatgtg 2421  
D 5878 TTGCTCTAGGGGGAGCCCGCCCGTGGCAGAAATGCTTGGCATAATGCTGATGTG 5937  
Q 2422 ctgcttcccgaccttgttctctagctcaagcgctccggtgtggtgtgaagcagcg 2481  
D 5938 CTGCTTCCACACCTTTGTTCTTATAGGCTCAAGGCGCTCGAGTGGTGAAGCAGACG 5997  
Q 2482 tgaagcagctcccggaagccttaactcagagagagactgtcaagttccacatcgcttcaa 2541  
D 5998 TGGAAGAGACCTCCGGAGGAGGCTACTCAGAGAGAGACTTGCAAGTTCCATCGCTTCAA 6057  
Q 2542 gttgaagatctgaactatgttagacgccaagctgcaagaacactcagagactcgttgaag 2601  
D 6058 GTTGAGAGATTCAGTATATGTATAGACGCCACCGGTGAGGAACCTCAGACATCGGTGAAG 6117  
Q 2602 ggaacttatctcgtactttgaaccaacaacagcgctgtgaaagtccgaagaaatcccaac 2661  
D 6118 GGACCTTATCTGTGTAATTTTGAACACACAGGCTGTGAAGATCGAAGGAATCCCAAC 6177  
Q 2662 tgaatcatalcatcccaagttlaagcgcgcacactcccgatctcggtgtgaagacgaa 2721  
D 6178 TGATTCATTCGATCCAGAGTTAAAGCGGCGGACACTCCGATTTGGGGTGGAAAGCCGAA 6237  
Q 2722 aagactgaaatccccttaagcttcgctcctacatcggtgtgttcttaactctgtcaataac 2781  
D 6238 AAGACTGAAATTCCTTAAACCTTCGCTCCATCGCGGTGTTCTTACTGTCTCAATAAC 6297  
Q 2782 tccctaaagttaagtlaagcctgtgtgagacgccaactccatataaccctatctc 2841  
D 6298 TCCTCAAGTTATATGTAAACGCTTGTGAGACGCCAGACTCCCATTAACCTTATCTCT 6357







QY 1282 gccgaagggaatcatalaacaattatagcagcaggtatgccttgcgacltgcacac 1341  
|||||  
Db 2294 gccgaagggaatcatalaacaattatagcagcaggtatgccttgcgacltgcacac 2353  
QY 1342 gfacaaggggccatctataaacaagggtgtgtcttactcaagcagggagaataaag 1401  
|||||  
Db 2354 gfacaaggggccatctataaacaagggtgtgtcttactcaagcagggagaataaag 2413  
QY 1402 aacaagaaggaaattcctaagcctatagaagccttactatctgccaanaagctatct 1461  
|||||  
Db 2414 AACCAAGAGGAAATCTTAAGCCTATTAGAGCCTTACATTTGGCCAAAAGGCTAGCTATT 2473  
QY 1462 atacactctctgagcatalcagaagaatcatalctatctagaagggaacagatgct 1521  
2474 ATACACTGCTCTGACATCAAGAACCCAAAGATCCCATATCCAGAGGAAACAGTGCT 2533  
QY 1522 gaccgggtgtcgaagcagcagcccaagctgttaacctctgcctctataatagaagccc 1561  
2534 GACCgggtgtcgaagcagcagcccaagctgttaacctctgcctctataatagaagccc 2593  
QY 1582 aaagcccaagaccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1641  
2594 AAAGCCCAAGACCAGGAG 2633  
QY 1642 gaccagctctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1701  
2654 GACCAGTCTCTGAGACATCCGGAAGGACCTGTATACCTCAGATGGGAAGGAATCCTG 2713  
QY 1702 ccccaagaagaagggttagaataatgctcaagaagatactgcttaaccactagaact 1761  
2714 CCCCAG 2773  
QY 1762 aaacacctgacagcagctgtgcaagaacatccctctatactgtctcgaagcagcagcagc 1821  
2774 AAACACCTGACAG 2833  
QY 1822 gctgactcgtgtgcaaacatctgtgctcctgcagcagcagcagcagcagcagcagc 1881  
2834 GCTGATTCGTGCTCAACACATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2893  
QY 1882 atactcagaagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1941  
2894 ATACTCCAG 2953  
QY 1942 actgaagtaagccggtcaaaatacgaagaacaatactatctgtcttctgaagcagcct 2001  
2954 ACTGAGGTAAAGCCGGCTAAATACGGAACAATATCTATGCTTTTGTAGACACTTT 3013  
QY 2002 tcaagatggttagagagcttactactaagaagaagacttcaaccgtgtgtgaagaa 2061  
3014 TCAGAGTGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCACCGTGGGTAAAGAA 3073  
QY 2062 atactgaggaatcttctcaagaattggaatacctaaggttaaggttaaggaatggt 2121  
3074 ATACTGAGGAAATTTTCCAAAGATTGGAATACCTAAGGTAATAGGCTAGCAAAATG 3133  
QY 2122 ccaagctcgtgtgccaagtaagtcaggaagctgtgccaagatcttgggagatgtatga 2181  
3134 CCAGCTTTCGTTGCCAGGTAAATCAGGAGCTGCCAAAGATTTGGGGATTGATGGAA 3193  
QY 2182 ctgcatgtgtacatacagaaccccaagctcagagcaggttagagagatgaatagaacat 2241  
3194 CTGATTTGTCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATAGAACATAT 3253  
QY 2242 aaagaagcccttacaatgacacagagacgtgcaataatgtgagctctctg 2301  
3254 AAAGAGACCTTACTAATTTGACCGGAGAGACTGGCTTATGATTGATAGCTCTCTG 3313  
QY 2302 cccttctgtcttcttgaagtgaggaacacccctgagcagcttggcgtgacccctatga 2361  
3314 CCTTTGTGCTTTTGAAGGTAGGAACACCCCTGACAGTTTGGGCTGACCCCTATGAA 3373  
QY 2362 ttgctctacagggagacccccccgtgtgacagaatctgcttgcacatagtgctgtgtg 2421

Db 3374 TTACTACGAGGGAGACCCCCCATTTGTGAATTTCTCCGACATATGCTGACGTG 3433  
QY 2422 ctgcttcccaagccttgtctcttagcctaaagcgctcgaagtgtgtgagcagcagc 2481  
3434 CTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCATTTAGTGAGGAGAACCAAGAGC 3493  
QY 2482 tgaagcagctccggagagcctactc--agagaggaacttgaagltccacatcttc 2538  
3494 TGGAGGCACTCCCGAGGCGCTACTCAGGAGGAGAGACTTGCAGATCCACATCTGTT 3553  
QY 2539 caagltgagatcagltatgttagacgcacacgtgacagaaacctcagagactcgtgtg 2598  
3554 CAACTGGAGATTCAGTCACTACGTAGAGAGGACGCGAGAGAAACCTCGAGACTCGTGG 3613  
QY 2599 aaggagcctatctcgtacttctgacacacacacacacacacacacacacacacacac 2658  
3614 AAGGCGCTTATACAGTATTTGACCAACCAAGGCTGTGAAGTTCGAAGGAATCTCC 3673  
QY 2659 acctgagctcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2718  
3674 ACCTGAGTCCATGATCCACAGCTTAAGCCGCGCCACCTCCGATTCGGGCTGGAAGCC 3733  
QY 2719 gaaagaactgaaatcccttgaagctcgcctccatcgcgtgtgtcttactcgtcgaat 2778  
3734 GAAAGACTGAATATCCCTTAAGCTTCCGCTCCTCATTCGGCTGCTTACTCTGTCAAT 3793  
QY 2779 aactcctcaagttatgtgtaaacgcctgtgtgacacccgaactcccaataacctatc 2838  
3794 AACTCTCAAGTTAATGTGAACGCTTGTGACACCCGAACTCCCATTAACCTCTTATC 2853  
QY 2839 tctcaactgttactactactcagcagcagcagcagcagcagcagcagcagcagcagcagc 2898  
3854 TCTCAGCTGGTTACTTACTACTCCGCTACAGGATTTAATATTAAACAGACTCAAGGGA 3913  
QY 2899 gactcccttggggagcctgtgtgctgaatatalgtctgcctcga tcaagtaacctgg 2958  
3914 GCTCCCTTGGGAGCGGAGGCTGAATTAATGCTGCTTCGATCAGTATCCCTGG 3973  
QY 2959 tctcaatgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3018  
3974 TCTCAATGACACAGGACACACCCCGATGATCTGCTGCTTACGCTTTTGGCC 4033  
QY 3019 aggacccccaataatgaagaatattgtgaaatctcagagcttctccttgcagaagcag 3078  
4034 AGGACCCCAATAATGAAGAAATATTGTGAAATCTCTCAGGATTTCTTTTGGCAAGCAATG 4093  
QY 3079 gaactgataacttcaatgataatggaatgtgaaatgtgcaagctcctcagcaagcaggt 3138  
4094 GAGCTCGTAACCTTAATGATGGAATGTGAAATGCGCAGTCTCAGCAGAGAGAGT 4153  
QY 3139 aagttactcttgttaacaatctcagcagcagcagcagcagcagcagcagcagcagcagc 3198  
4154 AAGTTACTCTTTGTTAAACATCTTACAGTTATATCAATTAATTAATGAGCCATGGAG 4213  
QY 3199 atggaagaatltgcaacacagcaggtacaaagaatgtgaaagaatgaagaatgaagctgtca 3258  
4214 ATGGAAGAATGTGCAACACGCGGTACAAAAGATGTACAAATTAACCAATTAAGCTGTCA 4273  
QY 3259 ttgcttagacactagattacttaaaataagtttcaataaaaaaataaaaaa 3317  
4274 TTTCGTTAGACCTAGATTAAATTAAGTTCACTGAAAGGAAAGCAAGAAATTA 4332  
Db

RESULT 5  
AX052634 7362 bp DNA PAT 12-JAN-2001  
LOCUS  
DEFINITION Sequence 19 from Patent W00071726.  
ACCESSION AX052634  
VERSION AX052634.1 GI:12226824  
KEYWORDS  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus

Viruses; Retrovirus; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.  
1 (bases 1 to 7362)

REFERENCE  
Fiederspiel, M.J.  
Methods to inhibit infectious agent transmission during  
xenotransplantation  
Patent: WO 0071726-A 19 30-NOV-2000;

JOURNAL  
Mayo Medical Ventures (US)  
Location/Qualifiers  
1. 7362

## FEATURES

source /organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

BASE COUNT 1997 a 1821 c 1881 g 1663 t  
ORIGIN

Query Match 95.7%; Score 3175.8; DB 6; Length 7362;  
Best Local Similarity 97.98; Pred. No. 0;  
Matches 3229; Conservative 0; Mismatches 67; Indels 3; Gaps 1;

QY 22 gatgctctctctgctgagatataccccaactagccaacacttttgcttcgaatg 81  
Db 3633 GATGCTCTTCTGCTGAGATTACACCCACTGACCAACACTTTTGCTTCGATGG 3692  
QY 82 agagatccaggtacggaagaacccggcagctcaactggaaccgactgcccgaaggtc 141  
Db 3693 AGAGATCCAGGTACGGGAAGAACCGGGCAGCTCAGTGGACCCGACCTGCCCCAAGGGTTC 3752  
QY 142 aagagatcccgaccatcttgagaaagccctacacaggaagcccgcccaacttcagatc 201  
Db 3753 AAGAGATCCCGACCATCTTGGAGAAAGCCCTACACAGAGACCTGGCCAACTTCAGGATC 3812  
QY 202 caaacacctcaggtgacctctcctcagtaagtgatgacctgcttcggggaggagccac 261  
Db 3813 CAACACCTCAGGTGACCTCTCTCAGTACGTGATGATGCTTCGGGGAGGCCACC 3872  
QY 262 aaacagagactgcttagaagtaagaaggaaggaaggaaggaaggaaggaaggaag 321  
Db 3873 AAACAGGACTGCTTAGAAGGACCAAGGACACTACTGCTGAAATGCTGACTGAGGCTAC 3932  
QY 322 aagacccctgctgtaagaagggccagattctgcaggaaggaaggaaggaaggaagga 381  
Db 3933 AAGACCTGCTGTAAGAAGGCCCAATTTGACAGAGAGGTAACATCTTGGGGTACAGT 3992  
QY 382 ttggcgaggcgagcagcagtgagctgacgaggaaggaaggaaggaaggaaggaagga 441  
Db 3993 TTGGCGGAGCGGAGCGATGCTGACGAGGACGGAAGAAACTGTGATGATCAATACCG 4052  
QY 442 gccccaacacagccaagaagtgagagagttttggggagacgtgatttgcagactg 501  
Db 4053 GCCCAACACAGCCCAAGAGAGAGTGTGTTGGGGAGAGCTGATGATTTGACACTG 4112  
QY 502 tggatcccgagggttgagcaacttagcagcccaactctacccgctaaccaagaagaagg 561  
Db 4113 TGGATCCCGGGGTTTGGCACTTAGCAACCCCACTTAACCACTTAACCACTTAACCACT 4172  
QY 562 ggaattctctgctgctcctgagcaccaggaaggaaggaaggaaggaaggaaggaag 621  
Db 4173 GAATTCCTCTGGGCTCCGAGACACCAAGAGCAATTTGATGCTATCAAAAAGGCCCTG 4232  
QY 622 aggcgactgctctgagccctcccgaggaacttaaacctttacccctttagtggatgag 681  
Db 4233 AGCCACACTGCTCTGGGCTCCGAGCTTAACCTTAACCTTTTATGTTGATGAGAG 4292  
QY 682 cgtlaagagagtagcccgagaggttttaacccaaccccttagagacatlgagaagagctgtc 741  
Db 4293 CGTAAGGAGTAGCCCGGGGAGTTTAAACCAACTCTAGACCATGAGAGAGAGCTGTT 4352  
QY 742 gctcaactgtaagaagaagcttgatcctgttagccagtggttggccgatagtcgtgaagct 801  
Db 4353 GCTCAACTGTCAAGAAGAACTCGATCTGTAGCCAGTGTGGCCGATGCTGAAAGCT 4412  
QY 802 atcgacagctgtgagcaactactgltcaagaagcgtgaacaaattgactttggagcagaata 861

Db 4413 ATGCGAGCTGTGGCATACTGTGTCAAGGACGCTGACAAATGACTTTGGACAGAAATATA 4472  
QY 862 actgtaaatgcccccatatgcatcttgagaaacatcgtctggcaagccccagacagatgagtg 921  
Db 4473 ACTGTAATGCCCCCCTCCATCGTTGGAGAAATCTGTTCCGACACCCCGACAGCATGAGATG 4532  
QY 922 accaagcccgatgaccacatacaagcctgtcttcaacagagaagggatcacttgcct 981  
Db 4533 ACCAAGCCCGCATGACCCACTATCAAGCCCTGCTTCCACAGAGAGGTCACGTTCCCT 4592  
QY 982 caacagccgctctcaacccctcctctctgcttggaggaagagatgagacagatgagct 1041  
Db 4593 CCACAGCCGCTCTCAACCTCCACTCTTGTGCTGAAGAGACATGATGAACCACTGACT 4652  
QY 1042 catgattgccaataactatgatttagagagagctgggttccgaaggaacttcaagacata 1101  
Db 4653 CATGATTGCCATCACTATTTGATTGAGAGAGCTGGGGTCCGAAGACCTTATACACATA 4712  
QY 4713 ccgctgactggaagaagtgctaaactggttactgaacggaagcagctatgtgtggaagct 1161  
Db 4773 AAGAGATGCTGGGGCGGGGCTGGTGGAGCGGACCCGACGATGGGGCAGACGCTTG 4832  
QY 1222 ccggaaggaacttcaacggaagaagctgagctcactgagccctacgcaagacttgcgctg 1281  
Db 4833 CCGGAGGAACCTTCAACCAAAAAGGCTGACGCTCATGGCCCTCCACCAAGCTTTGGGGCTG 4892  
QY 1282 gccgaaggaagaatccaataacttatacagcaaggaagtaagcttgcgactgcaac 1341  
Db 4893 GCCGAGGAGAAATCCATTAACATTTATACAGACACAGGATTAAGCCCTTGGCGCTCACAC 4952  
QY 1342 gtaacaggggcatcctataaacaagaagggttgccttaccctgaacaggaaggaataag 1401  
Db 4953 GTACAGGGGCGCACTCATTAAGCAAAAGGGTGTGCTTACTTCAAGCAGAGGAAATTAAG 5012  
QY 1402 aacaagaagaagaatcttaagcctatagaagccttaccatcttgcgaagaagctagctat 1461  
Db 5013 AACAAAGAGAAATTTGTAAGCCATTAGAAAGCTTACATTTGCCAAAAGGCTTACTATTT 5072  
QY 1462 atagactgtcctgagacatcaagaagccaagaatctcatalctcaagggaggaacagtgct 1521  
Db 5073 ATACACTGTCTCGACATCAGAAAGCCAAAGATCCCATCTTCCAGGGAGAACAGATGCT 5132  
QY 1522 gaccgggttgccaagagagcagccagagctgtttaaactctgctcctataatagaagagcc 1581  
Db 5133 GACGGGGTGGCCAGAGAGGACGACCGGCTGTTAACTTCTGCTATGATTAACACACC 5192  
QY 1582 aaagccccaagaacccaagaagtagacacccctagaagactggaagagaataaagaata 1641  
Db 5193 AAAGCCCCAAGAACCCGAGACAGATGACCCCTAGAAAGACTGCGCAAGAGATTAAGAAATA 5252  
QY 1642 gaccagttctctgagactcccggaaggagactcttaactctataaggaagaagaatccg 1701  
Db 5253 GACCAATTTCTGAGACTCTCGAAGGAGCTCTATACCTCGAATGGGAGAGAAATCCG 5312  
QY 1702 ccccaagaagaaggttgaataatgctcacaagataacatcgtcttaacccaacccaaggaact 1761  
Db 5313 CCCCAAGAAAGAGGTTAAGATATGTCACAGATATCATGCTTAACCCACTAGAACT 5372  
QY 1762 aaacactgcaagcagttgltcagaacatccctcatalcaltgltcgaagcttacccagagtg 1821  
Db 5373 AACACCTGACAGAGTGTGTCAGAACATCTCTTATCATGTTCTGAGGCTTACAGAGAGTG 5432  
QY 1822 gctgactggtgtgtaacaatgltggtgcccctgcagcgtgtgtaagcttaactctccaga 1881  
Db 5433 GCTGATTCGCTGCTAAACATGCTGTGCTCCAGCTGCTGTTAATGCTAATCTTCCAA 5492  
QY 1882 ataccctcaggaagaagactaaagggaagccacccagcgctcactggaagtgagcttc 1941

Db 5493 ATACCTCCAGGAAAGAGACTAAGGGAGAACCCACCGCTCCTACTGGAGATGACATTC 5552  
 QY 1942 actgaggtaaagccggtaataatagaaacaaatctctatgttttttgaagcactt 2001  
 Db 5553 ACTAGGTTAAAGCCGGCTTAATACGGAAACAAATATCTATGTTTGTAGACACCTTT 5612  
 QY 2002 tcaagatgggtagaggtctactactaagaagaagactcaacccgtgtgtgctaaagaa 2061  
 Db 5613 TCAGGATGGGTAGAGGTTATCTACTAAGAAAGAGACTTCAACCGTGGTGTAAAGAA 5672  
 QY 2062 atactgaggaataatlttccaagatttgaaatccaaagtaagtgatgaagaaatgt 2121  
 Db 5673 ATACTGGAGGAATTTTCCAAATTTGGAATTTGAAATAGGTATAGGTGACAAATGGT 5732  
 QY 2122 cgaagcttggtcccaaggaatgcaggactggtgccaagatattgggatttgaa 2181  
 Db 5733 CCAAGCTTTCGTGCCCGGTAAGTCAAGGACTGCGCAAGATATTGGGGATTGTTGGAAA 5792  
 QY 2182 ctgcatgtgcatagacagcccaagactcagagcaggtagagagatgatatgaacatt 2241  
 Db 5793 CTGCATTTGTSCATACAGAACCCCAAGCTCAGGACAGTGAAGATGAAATGACACATT 5852  
 QY 2242 aaagagacccctaccaaattgaccacagagactgcatatgattgattggtctctg 2301  
 Db 5853 AAAGAGACCCCTTACTAATTTGACGGGAGACTGGCGTTAATGATTGATAGTCTCTG 5912  
 QY 2302 ccccttggtgtttttaggggtgagaaacacccctgacattgggtgtgagccctatgaa 2361  
 Db 5913 CCTTTGTACTTTTATAGGGTTAGGAACACCCCTGACAGTTGGGTGACCCCTATGTA 5972  
 QY 2362 ttgcttaagggggagcccccctgtgtgcagaatctgcttgcacatgtgtgtgt 2421  
 Db 5973 TTACTCTACGGGGAGACCCCGCCCATTTGTAGAAATTTGTTCCGTGATAGTCTGACGT 6032  
 QY 2422 ctgcttccagacttgttcttcttaagcccaagcgctcgatgggtgagcagcgagc 2481  
 Db 6033 CTGCTTTCCAGGCTTTGTTCTCTTACGGCTCAAGGCACTTGAAGTGGTGAACACGAGCG 6092  
 QY 2482 tgaagcagactccgggagagcctactc---aggagagacttcaagttccacatgcctc 2538  
 Db 6093 TGGAGGCACTCCGGGAGGCTCACTCAGAGAGAGAGACTTGCAGATCCCAATCGTTTC 6152  
 QY 2539 caagttggagatcagctctatgtttagacgcacccgtgcagagaacctgcagactcgt 2598  
 Db 6153 CAAGTGGAGATTCAGTCTACGTTTACGCCCAACCGCTGAGGAACCTCGAGACTGGTGTG 6212  
 QY 2599 aagggaccttatctgacttcttgaccacacaaagcgctgtgaaagtcgaagaaatccc 2658  
 Db 6213 AAGGGCCCTTATCACGTACTTTTGGACACCAACGCGCTGTGAAGTCAAGAAATCTTC 6272  
 QY 2659 acctgcatcagcatcccaagcttgaagcgcgccacccctccgattcggtgtgaaagcc 2718  
 Db 6273 ACCTGATTCATGATCCCAAGCTTAAAGCGCGGCCACCTCCGATTCGGGGTGAAGGCC 6332  
 QY 2719 gaaaagactgaaatcccttaagctctgcctccacacgcgtgtgttccctactctgcaat 2778  
 Db 6333 GAAAAGACTGAAATCCCTTAAAGCTTCGCTTCATCGCGGTGTTCTTACTCTGTCAAT 6392  
 QY 2779 aactctcaagttaagtgaagcgctgtgaaagcccggaactcccaataaaccctatc 2838  
 Db 6393 AACTCTCAAGTTAATGTTAAAGCGCTTGTGACAGCCCGGAATCCCAATAACCTTTATC 6452  
 QY 2839 tctcacttgattactactgactcagtcagatlaattlaattacagaaactcaagggga 2898  
 Db 6453 TCTCACCTTGTTACTTACGACTCCGGTACAGSTATTATTAACACGACCTCAAGGGGA 6512  
 QY 2899 ggtcccttgggaacctgtgtgcgtgaattatattgtcgtccttcagtaacatccctg 2958  
 Db 6513 GGTCCCTTGGGGACCTGTTGGCTGAATTAATATGCTCTTCATCGATGATATCCCTGG 6572  
 QY 2959 tctcaatgacagggacacaccccgatgatactcgtgtcttaaggttttaagttgcc 3018  
 Db 6573 TCTCAATGACAGGCGACACCCCGGATGACTCGTGTCTTACGGGTTTACGTTTGGCC 6632

QY 3019 aggaccccccaataatgaagaatattgtggaatccctagatttcttgcagcaatg 3078  
 Db 6633 AGGACCCCAATTAATGAAGAAATATTGTGAATCTCAGATTTCTTTCCAGCAATG 6692  
 QY 3079 gagctgcaatcttcaatgaatgaggaatcggaaatggcagctctcagcaagcagat 3138  
 Db 6693 GAGCTGCGTAACCTTCTAATGATGGGAATTTGGAATGGCCAGTCTTCAGCAAGACAGT 6752  
 QY 3139 aagttactcttcttgaacaaatcccaagttataatcaatlaattlaattgagcaatgag 3198  
 Db 6753 AAGTACTCTTTGTTAAACAATCTACAGATTAATCAATTAATTAATGAGCAATGGAG 6812  
 QY 3199 atggaagattggtcaacgcggtgtcaaaaagatgtacgaataatgaacaaatgctca 3258  
 Db 6813 ATGGAAGATTGTGGCAACGCGGTCAAAAAGATGTACGAATAATGAAATAGCTGTCA 6872  
 QY 3259 ttgcttagacctaagttacttaaaataagtttcaactaaataaaataaaataaa 3317  
 Db 6873 TTGCTTAGACCTAGATTACTTAAATATAGTTTTCACCTGAAAAAGGAACAAAGAAATA 6931  
  
 RESULT 6  
 AX052638  
 LOCUS AX052638 7873 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 23 from Patent WO0071726.  
 ACCESSION AX052638  
 VERSION AX052638.1 GI:12226828  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 porcine endogenous retrovirus.  
 porcine endogenous retrovirus  
 Viruses; Retroid viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 REFERENCE  
 AUTHORS  
 Federpspiel, M.J.  
 TITLE  
 Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 JOURNAL  
 Patent: WO 0071726-A 23 30-NOV-2000;  
 Mayo Medical Ventures (US)  
 FEATURES  
 source  
 1. 7873  
 Location/Qualifiers  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 2200 a 1914 c 1961 g 1798 t  
 ORIGIN  
 Query Match 89.7%; Score 2978; DB 6; Length 7873;  
 Best Local Similarity 95.2%; Pred. No. 0;  
 Matches 3140; Conservative 0; Mismatches 70; Indels 90; Gaps 3;  
  
 QY 22 gatgcctctctgcgtgagatcaacccacatagcaaacacttctgctcgaaatg 81  
 Db 2447 GATGCTTTTCTGCTGAGATTACACCCCACTAGCCCAACACACTTTTGGCTTGAAATG 2506  
 QY 82 agagatcaagtagaggaagacccggcagctacccctggaccgactgcccgaaggttc 141  
 Db 2507 AGAGATCCAGGTACGGGAGAGAACCGGAGCTCACCTGAGACCGAGCTGCCCAAGGGGTTCC 2566  
 QY 142 aagaactcccgagacatcttgaagaaagccctacacagggagacccgagcaacttcaagatc 201  
 Db 2567 AAGAAGCTCCCGACACATCTTTGAGAGAGCCCTTACAGAGACCTGGGCAACTTACGATC 2626  
 QY 202 caaacacctgaagtgaacctctcctcagtaagtgatgacctgcttctggcgaggacacc 261  
 Db 2627 CAACACCTCAGGTGACCTCTCTCAGTACTGATGACTGCTTCTGGCGGGAGCCACC 2686  
 QY 262 aaacagactcttagaaggtgaggaagcactactgctgaatgtgtgagcaggtac 321  
 Db 2687 AAACAGACTGCTTAGAAAGACGACAGGCACTACTGCTGGAATGTCTGACTGAGCTTAC 2746  
 QY 322 agagcctctgtaagaagggcccaatcttgacaggaaggttaacatacttgggtacagt 381

DB 2747 AGAGCTCTGCTAGAGAGCCCAATTTGACGAGAGAGGTAACTACTTGGGTACACT 2806  
QY 382 ttgcgggagcgggagcagctgagctgacgagagcaggaagaaactgtatgacagatacgc 441  
DB 2807 TTGCGGAGCGGAGCGATGCTGACGAGGACGAGGAAGAACTGTATGTCAGATACCG 2866  
QY 442 gccccaccacagcagcaagctgagagagcttttgggagagctgtatgttcagactg 501  
DB 2867 GCGCCAAACCAAGCCAAACAGAGAGATTTTGGGACAGCTGATTTTGGAGACTG 2926  
QY 502 tggatcccgagggtttgacagcttagcagcccccactctacccgctcaacaaagaaaggg 561  
DB 2927 TGGATCCCGGGGTTTGGACACTTGAAGACCCCACTATACCACTAACCAAGAAAGAGG 2966  
QY 562 ggaattctcctgagctctgagcaccagagagcatttgaatgatalcaaaagccctgctg 621  
DB 2987 GAATTTCTCTGGGCTCTGAGCACCAGAGCATTTGATGATCAAAAGAGCCCTGCTG 3046  
QY 622 agcgacactgctctgagcctccctcgaagtaaacaccttaccctttatgtgagag 681  
DB 3047 AGCGACCTGCTCTGGCCCTCTGAGCTAACTAAACCTTTTACCTTTATGTGATGAG 3106  
QY 682 cgtaaagagtagcccgagagagtttaacccaacccctagagacatgagagagccctgtc 741  
DB 3107 CGTAAGGAGTAGCCCGGGGAGTTTAAACCAACTAGAGACCATGAGAGAGACTGTT 3166  
QY 742 gctacactgtcaaaagagctgtatcctgttagccagtggttggccgtatgtcgaagct 801  
DB 3167 GCTTACCTGTCTAAAGAGCTGTATCTGATCCAGTGTGGCCGCTGATGCTGAAGGCT 3226  
QY 802 atcgacactgtgagcactgtgtcaagagcgtgacaatgtgactttggagagagata 861  
DB 3227 ATTCGAGCTGTGGCCATTAAGTGTCAAGAGCCTGACAAATGACTTTGGAGAGAAATA 3286  
QY 862 actgtaagccccccatgcatlgtgagagacatgttcgagagcccccagacagatgagt 921  
DB 3287 ACTGTAAATAGCCCCCACTGAGCTGTGAGACAACTGCTGCGAGCCCAAGACATGATG 3346  
QY 922 accaagcccgcatgagcccaactatcaaaagctgtcttcaagagagagtgcaacttgcgt 981  
DB 3347 ACCAAGCCCGCATGAGCCCACTATCAAAAGCTGCTTCTCAAGAGAGGTCAGCTTGCT 3406  
QY 982 ccacacgacgctcacaacctgtccactctctgctgaagagactgtatgacagagtagt 1041  
DB 3407 CCACAGCGGCTCTCAACCTGCGACCTTCTGCTGAAGAGACTGTATGAACAGTACT 3466  
QY 1042 catgattgccaactaatgtatgtgagagagctgggttcgcgaagagacctaagacata 1101  
DB 3467 CATGATTGCCATCAACTATTGATTGAGAGAGACTGGGTCGCAAGACCTTACAGACATA 3526  
QY 1102 ccgctgactgagagagagctgcaactgtgtactgagagcagagcagctatgtgtgagaggt 1161  
DB 3527 CCGCTGACTGAGAGAGTGTAACTGCTGTTCACTGAGGAGAGCATGTATGTAGAGAGGT 3586  
QY 1162 aagaagatgtgctgagcagtggtgtgagcagagaccccgacagatctgtgagcagagcgt 1221  
DB 3587 AAGAGATGTGGTGGGGGGGGGTGTGAGAGGAGCCGACAGATCTGGGCGACAGCTG 3646  
QY 1222 ccggaagagactcagcgcgaagagctgagctatgagccctcaagcagagcttgcggtg 1281  
DB 3647 CCGGAAGGAACTTCAGCAAAAGGCTGAGCTATGCTCCTCAAGCAAGCTTTGGGCTG 3706  
QY 1282 gcccgaagagaaatccataaactttatagagagagcagagatgacttggactgagacac 1341  
DB 3707 GCCCAAGGGAATTCATTAACATTTATACAGACAGGAGTATGCTTTGGACTGACAC 3766  
QY 1342 gtacacggggccatctataaacaagagggtgtgcttaccctcagcagaggaagaaataag 1401  
DB 3767 GTACACGGGGCATCTATAAGCAAAAGGGGTGTGTTACTACAGAGGAGAGAAATAAG 3826  
QY 1402 aacaaagagaaatctcaagccatctagagagccttaccattgccaagagcagctatgt 1461  
DB 3827 AACAAAGAGAAATTTTAAAGCTATTAGAGCCTTACATTTGCCAAAAAGGCTAGCTATT 3886

QY 1462 atacactgtcttggacatcagaaagccaaagatctcatctagagaggaaccagatgct 1521  
DB 3887 ATACACTGTCTGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAAACAGATGGCT 3946  
QY 1522 gaccgggttgcagagcagcagcccgctgttacccttctcctataatagaagccgc 1581  
DB 3947 GACCGGTTGCCAAGAGAGGAGCCCGAGGGTGTAAACCTTCTGCTATGTATGAATACACC 4006  
QY 1582 aaagcccccagaaacccagacagtagtacacctagaaagactgtgcaagagataaagata 1641  
DB 4007 AAAGCCCAAGAACCCGAGAGCAGAGTACACCTTGAAGAGACTGGCAGAGATTAAGAAAGATA 4066  
QY 1642 gaccagttcttgaagactcccgagagagcctgtatacctcatatggaagaaatccgt 1701  
DB 4067 GACCAATTCCTGAGACTCCGGAAGGAGACTGCTATACCTCAGATGGGAAGAAATTCCTG 4126  
QY 1702 ccccaaaagaaaggttagaataatgccaacagatcatcgtcttaacccacttaagaact 1761  
DB 4127 CCCCAAAAGAAAGGTTAAGATATGTCAACAGATACATGCTTAACCCACTAGAACT 4186  
QY 1762 aaacacctgcagagctgtgtcagaacatcccttatcatgttctgaagcttacagagtg 1821  
DB 4187 AAACACTGACAGAGTGTGTCAAGACATCTCTTATCATGTTCTGAGGCTACAGAGATG 4246  
QY 1822 gctgactcgtgtgcaaacatgtgtgcccgtgcagctgtgttaatgtcaatcctccaga 1881  
DB 4247 GCTGATTCGTTGTCAAAACACTGTGTCCTGCTCCAGCTG----- 4285  
QY 1882 ataccctcagaaagagactaagagggaagcaaccagagcctcactgtggaagtgaatc 1941  
DB 4286 ----- 4285  
QY 1942 actgaagtaagccgagctaaataacgaaacaaatctatgtgttttgaagacactt 2001  
DB 4286 -----GCTAAAGCCGCTTAATACGSAACAAATATCTATGTGTTTGTATACACCTTT 4340  
QY 2002 tcaagatgtgtgagagcttatacctactaaagaaagagactcaacggtgtgtcagaag 2061  
DB 4341 TCAGATGTGGTGTAGAGCTTATCTCTTAAGAAAGAGACTCAACCTGCTGCTTAAGAA 4400  
QY 2062 atactgga-ggaatatttccaaagatttggaaatacctaagagtaatgaggttcagaagatg 2120  
DB 4401 ATACTGAGAGGAATTTTTCAGATTTTGAGATTAAGATTAAGGCTGAGCAATAGG 4460  
QY 2121 tccagcttgcgtgccagagtaagtcagagagctgtgccaagatatgtggatgtatgaa 2180  
DB 4461 TCCAGCTTTCGTTGGCCAGGTAAAGTACAGGAGCTGCGCAAGATATGGGATTTGGA 4520  
QY 2181 actgacatgtgtatatacagcccccaagctcagagaaagttaagagagatgaataagacat 2240  
DB 4521 ACTGATTTGTGATACAGACCCCAAGCTCAAGAGAGTGAAGAGATGAATGAACAT 4580  
QY 2241 taaagagaccccttacaataatagacacagagagctgtgcaataatgtatgtgagcctct 2300  
DB 4581 TAAAGAGACCTTACTTAATTTAGCCGAGAGACTGGCTTATGTATGTGATAGCTCTCT 4640  
QY 2301 gcccttgtgtcttlaaggtgaagaaacacccctgtgagagcttggcgtgacccctatga 2360  
DB 4641 GCCCTTGTGCTTTTATGAGTTAGGAACACCCCTGACAGTTGGGCTGACCCCTATAA 4700  
QY 2361 attgctctcaggggagaccccccgtgtgacgaatgtccttggacataatgtcgtatgt 2420  
DB 4701 ATTACTCTAAGGGGAGCCCCCTCTGTTGTAATAATTCCTTCCATCAATGCTGAGAGT 4760  
QY 2421 gctgtcttccagccttgttctcagagctcaagcgtgtgagtgagagcagcagc 2480  
DB 4761 GCTGCTTCCAGCCTTTGTCTTCTAGGCTCAAGGACTTGTAGTGGGTGAGACAAAGAGC 4820  
QY 2481 gtggaagcagctcggagagcctactc---aggagagagcttgcagatltcacatgctt 2537  
DB 4821 GTGAGAGCAACTCCGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCCAATCTGCTTT 4880

QY	2538	ccggttgaggaatttaagctatcgtatgatagaacgcacacggctgaaggaaacctcgaaactcggty	2597
Db	4881	CCAAGTGGGATTTCAGCTACGTATGAGCCACCCTGGAGGAACCTCGAAGACTCGGtG	4940
QY	2598	gaagagaccatactcgtacttttgcacacacacacgagctgtgaaagtlogaagyaatccc	2657
Db	4941	GAAGAGGCCCTTATCAGGTACTTTTGACCAACACAAAGCGCTGTGAAGTGAAGAATCTC	5000
QY	2658	caactgatalcatgcalcccaagttaagccgycacaactcccgatcggggtgaaaagc	2717
Db	5001	CACCTGGATCCATGCATCCACAGCTTAAGCCGGGCGCACCTCCGATTCGGGGGTGAAGAGC	5066
QY	2718	cgaagaagacccaatatccccctaagtttagtcgccatcgcacgcgfygtttccctactcgtcaa	2777
Db	5061	CGAAAAACAAGAAAATCCCCTTAGCTGTGCCTCCATCGCGTGGTTCCTTACTCTGTCAA	5120
QY	2778	laactcccaagttaatgylgaaacagcgcctgtgtgacagcccgcaactoccataaaccttat	2833
Db	5121	TAACTCTCAAGTTAATGATAACGGCTTGTSACAGCCCCGAATCCCATTAACCTTAT	5180
QY	2838	ctctcacctggttactactgactgactccgylgacaagattaataaltaacagcactcaaggg	2897
Db	5181	CTCTCACTTGTTACTTACTACTCGCGTAAAGGTATTAAATTAACACACTCAAGGGG	5240
QY	2898	aagctcccttggygagactgtytgccggaattatgltcgcgcttgatagaatccccy	2955
Db	5241	AGGCTCCCTTGGGAGACTGGTGGCTGGAATTATATGTGCTTCGATCAGTAATCCCTG	5300
QY	2958	gtctcaatgacacagccacaccccccgatgctactccgtytcttaagvggttlcagttlcc	3017
Db	5301	GTCATCATGACAGGCCACACCCCCCGATGTACTCCGTGTACGGGTTTTACGTTTGC	5366
QY	3018	caggacccccnaaatgaaatlatltgylgaaatccctcaagatlctccttgcgcaagcaat	3077
Db	5361	CAGAGCCCCCAATAATGAGAATATTGTGGAATCCTCAGGATTTCTTTTSCAAGCAAT	5420
QY	3078	ggagctgcataactctcaatgaglygggaatttggaaatggccagfctctcagaagacagag	3137
Db	5421	GGAGCTCGTPAACCTTCAATGATGGGAGATTGAAAAGGCCAGCTCTTCAGCAAGACAGAG	5480
QY	3138	taagttaactcttltgttaacaatccctacacagttatatcaatttaatttaattgacatygga	3197
Db	5481	TAACTTACTCTTTTGTTAACAAATCCATACGATTAATCAATTTAATTATGGCCATGSA	5544
QY	3198	galgaaaaagattlgcaacacgcyggtlacaaaaagatglacgaaataagcaataagctgtlc	3257
Db	5541	GATGGAAGAAGATTGGCACACGCGGTACAAAAAGATGTAGCAAAATAGCAATAAGCTGC	5600
QY	3258	attcgcttagacctagatlaactttaaaaaataagtttcaacttaaaaaaaaaaaaaaaaaa	3317
Db	5601	ATTGCTTAGACCTTAGATTACTTTAAAAATAAGTTTCACCTGAAAAAGGAAAAACAAGAAATA	5666
RESULT	7		
LOCUS	A66552	8196 bp DNA	PAT 29-MAR-1999
DEFINITION	Sequence 2 from Patent WO9740167.		
ACCESSION	A66552		
VERSION	A66552.1 GI:4538105		
KEYWORDS			
SOURCE			
ORGANISM			
	porcine endogenous retrovirus. porcine endogenous retrovirus viruses; Retrod viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group. 1 (bases 1 to 8196)		
REFERENCE			
AUTHORS	Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.		
TITLE	PORCINE RETROVIRUS		
JOURNAL	Patent: WO 9740167-A 2 30-OCT-1997;		
FEATURES	Q ONE BIOTECH LTD (GB) location/Qualifiers 1..8196		
SOURCE	/organism="porcine endogenous retrovirus" /db_xref="taxon:61673"		

[illegible]

Query Match	87.9%;	Score 2916.8;	DB 6;	Length 8196;
Best Local Similarity	96.0%;	Pred. No. 0;		
Matches 3013; Conservative	2;	Mismatches 114;	Indels 9;	Gaps 2;

OY	19	acagatgcctctctcttgcgcctgagatttaacccaacacatagccaacacacttttgccttcgaa	78
Db	2956	AAATGTCCTTTCTTTCGCTCGAATATTACACCCACAGCAACCACTTTTGGCCTTGSA	3015
OY	79	tgcagagatccagttacggygaagaacccgygcagcttaacctgtaaccgcagctgcccaagg	138
Db	3016	TGGAGAGATCCAGGTACGGGAGAAACCCGGGAGCTCACCCTGAGCCCGACTGCCCAAGG	3075
OY	139	ttcaagagactcccgagacatcctttgacgaagccctcaacaaaggagacttggccaactcaag	198
Db	3076	TTCAAGAACTCCCGGACCATTCTTTGAGAGAGCCCTTACACAGGAGACTTGCCAACTTCAGG	3135
OY	199	atccaaacacctcaagctgcagccctccaccagtaagttgcagctcttcttgcggyagcc	258
Db	3136	ATCCAAACCCCTCAGGTGAGACCTCTCTGCACTGAGTGGAGACCTGCTCTGCGGGAGCC	3195
OY	259	accaaaacagagctgtcttaagaagtgacgaagcaactacgcctgcyyaatgtctgcacctagc	318
Db	3196	ACCAAAACAGAGCTCTTATGAAAGGTACGAAGGCACTACTCTCGAATTGTCTGCACCTAGCG	3255
OY	319	tacagagcctctgtttaagaagcccgagattttgagagaagagttacactttgggtgac	378
Db	3256	TACAGAGCTCTGTCTTAAGAAAGGCCCAATTTGAGAGAGAGGTATACATCTTGGGGTAC	3315
OY	379	agttgcggygcggygcagatgtgcctgcagcgagcgacggaagaanaacttgatccaaga	438
Db	3316	AGTTTGGGGGGCGGGGCGAGCATGGCTGAGCGAGGACGGAAGAAAATGTATGTCAGATA	3375
OY	439	ccggcccacaacacagccaagaatgagagagtttttggggaacagcttggattttgcaga	498
Db	3376	CCGGCCCCAACACAGCCAAACAAAGTAGAGATTTTGGGAGACAGCTGGATTTTGCGAGA	3435
OY	499	ctgctgagatcccggyggttgcgaccttagacagccccaacttacccgtacaaccaagaanaa	558
Db	3436	CTGCTGATCCCGGGGTTTGGGAACTTACGACGCCCACTTACCCGCTAACCAAGAAAAA	3495
OY	559	ggggagatctcctctgccttccctgcagcaacgaagaagcatttgcattgcatacaaaaagccctg	618
Db	3496	GGGGATTTCTCGTGCGCTCTGAGACACAGAAAGCATTTGATGCTATCAAAAAGGCCCTG	3555
OY	619	ctgagcgacactgctctgcgcctccctgcagctaaacttaaccccttaacttattgtgat	678
Db	3556	CTGACCGCACTGCTCTGCGCCCTCCCTGCACTAACTAAACCCCTTATACCTTTATGTGAT	3615
OY	679	gagcgttaagagagtagcccgagagagtttbaaccaaacctttaggacaaatgcggagagact	738
Db	3616	GAGCCTAAGGAGATGACCCCAAGAGTTTATACCAACCCCTTAGACATGAGAGAGAGACT	3675
OY	739	gttgcctlaactgcgtcaagaagacttgatccctgttagccaagtggttggccgtatgctcgtgaag	798
Db	3676	GTTGCTACTGCTCAAGAAGACTTGATGCTGTAGCCAGTGTTGGCCGCTATGCTGTGAAG	3735
OY	799	gctatcgagctgtgcccatactgctgcacaagaagcgtgcacaattgactttggagaagat	858
Db	3736	GCTATCGAGCTGTGGCCATTACTGTGTCAAAGAGCGCTGACAAATTGACTTTGGGACGAT	3795
OY	859	ataacttgaatagccccccatgcatactggaanaacatgcttcggcaagcccccaacgaatg	918
Db	3796	ATAACTGTAATAGCCCCCATGCACTTGGAGAGACATGTTTCGGCAGCCCCCAAGCCGATGG	3855
OY	919	atgacaacagccgcgatgcacacatacaagaacgtgcttctcaagagagaaggtcaattc	978
Db	3856	ATGACCAACGCCCGCATGACCACCACTATCAAAAGCTGCTCTTCACAGAGAGGGTCACTTTC	3915
OY	979	gtctcacacagcgcctctcaaacctgtcacactctctgcctcgaagagatcgtatgaacagtg	1038

|||||  
Db 3916 GCTCCACGACCGCTCTCAACCTGCGACCTTTGCGCCGGAAGACATGATTAACCGAGTG 3975  
OY 1039 actcatgattgcatcaatcatattgattgaggaactggtccgcaagaccttacaag 1098  
Db 3976 ACTCATGATTGCCATCACTATTGATTGAGGAGACTGGGGTCCGCAAGACCTTACAGAC 4035  
OY 1099 ataccgtactgaggaagtgctaaccttggttcaactgacggaagcaatgltggtgaa 1158  
Db 4036 ATACCCCTGACTGAGAAATGCTTAACCTGTTCACTGACGGAAGCAGCATATGTGTGMA 4095  
OY 1159 gttaaagagatgctggtggcgagtggtgagcgagaccgagatctggcgagcgc 1218  
Db 4096 GGTAAAGAGATGCTGGGGCGAGTGGTGAACGGGACCCGACGATCTGGGCGACGACG 4155  
OY 1219 ctgcccgaaggaactcagcgcaaaagctgagctcatgagccctcaagcaagcttgcg 1278  
Db 4156 CTGCGGAGGAACTTACACCGCAAAAGGTGAGCTCATGCGCCCTCACCGCAAGCTTTGG 4215  
OY 1279 ctgcccgaaggaatcctaataacattatagcaagcaggtatgcttctgagctgca 1338  
Db 4216 CTGGCGGAAGGGAATCCATAACATTATACGACAGGATGCTTGGCCACTGCA 4275  
OY 1339 caagtacacggggccatctataaacaagggtgtgcttactcaagcgagggagaaata 1398  
Db 4276 CACGTACACGGGGCCATCTATTAACAAGGGGTTGCTTACCTCAGCAGGAGGAATAVA 4335  
OY 1399 aagaacaagaaggaatcttaagcctatagaagccttacaatttgcacaaagctagct 1458  
Db 4336 AAGAACAAGAGAGAAATCTTAAGCTATTAGAGGCTTACATTTGCCAAAAGGCTACT 4395  
OY 1459 attatacactgctcgtgagacatcagaagccaaagatcctaatactagaaggaaacag 1518  
Db 4396 ATTATATACACTGCTCTGGACATCCGAAAGCCAAAGATCTCATATCTAGAGGGAACCA 4455  
OY 1519 gctgacgggtgttgcgaagcagcgccagcggtgttaaccttctgacctataatagaag 1578  
Db 4456 GCTGACCGGGTTCGCCACGAGGACGACCCAGGCTGTTAACCTTCTGCTTATATATGAAC 4515  
OY 1579 eccaaagcccagaacccaagcagcaglacacccctagaagaactggtcaagaagataaaag 1638  
Db 4516 CCCAAAGCCCGCAGAACCCAGACAGACAGTACACCTAGAAAGCTGGCAAGACATTAAGA 4575  
OY 1639 atagacagttctctggaacatccgagagagactgctataactcataatggaagaaatc 1698  
Db 4576 ATAGACAGTCTCTGGAACCTCCGAGGGGACCTGCTATACCTATATAGGGAAGAAATC 4635  
OY 1699 ctgcccacaagaagggttagaatatgtccaacagatatacgtcttaaacccactaga 1758  
Db 4636 CTGCCCAACAAGAGGGTTAGAAATATGTCCACACAGATACATGCTTAACCCACTAGGA 4695  
OY 1759 actaaacactgagcagctgtgtcagaacatcccttacaatglttctgaggtaccaga 1818  
Db 4696 ACTAAACACCTGCACACCACTGGTCAAGAACATCCCTTATCATGTGTTGAGGCTACCA 4755  
OY 1819 gtggtgactcggtgtgtaaacattgtgtgcccgcgaagcggttaagtaacatccctc 1878  
Db 4756 GTGGCTGACTGCTGTAACCATTTGTTGCTGCGACGCTGTTAAATGCTTAATCTTCC 4815  
OY 1879 agaatacctcagaagaagactaaggaaggaacacccagcgctcactgtggaagtgtac 1938  
Db 4816 AGAATACCTCCAGGAAGAGACTAAGGGAAGCCACCCAGGCGCTCATGTGGGAAGTGGAC 4875  
OY 1939 ttcaactgagtaagcgggttaaatcaggaacaataatctatgtgttttgtgagacac 1998  
Db 4876 TTCACTGAGTAAAGCCCGCTAAATACGGAACAAATATCTATTGTTTGTGACACCC 4935  
OY 1999 ttctcaggaatgtagaggtctatacctaactagaagaagacttcaacccgtgtgtcgaag 2058  
Db 4936 TTTCAGGATGGTAGAGCTTATCTACTTAAGAAAGAGACTTCAACCCGTCGTGCTAG 4995  
OY 2059 aaatactgaggaatcttccaagattggaatacctaaggaataggtcagaat 2118  
|||||

---

Db 4996 AAAATACTGAGAAATTTTCCAAAGATTGGAAATACCTAAGTAATAGGTCAGACAT 5055  
OY 2119 gtccagacttcgttgcacaaagtaagtaaggaactggtccaaagattggtgattgtg 2178  
Db 5056 GGTCCAGCTTTCGTTGGCCAGGTAGTCAAGGACTGGCCCAAGATATTGGGATGTGG 5115  
OY 2179 aaactgacttgatatacagaaccccaagctcaggaagaagtaagagaatgaatgaac 2238  
Db 5116 AAATCGATTTGCTATACAGACCCCAAGAGCTCAGAGACAGGTAGAGAGATTAATGAAAC 5175  
OY 2239 ataaagaagcccttaccatgaacacagagactgcatatgattgattgtgtctc 2298  
Db 5176 ATTAAGAAGACCTTACCAATTAAGCACAGAGACTGGCAATATGATTGATGGCTCTC 5335  
OY 2299 ctgccccttgtgcttcttgaagtgaggaacaccccttgaagaattggtggtgacctat 2358  
Db 5236 CTGCCCTTTGCTTTTAAAGGTAGGAACACCCCTGACAGTTTGGGCTGACCCCTAT 5395  
OY 2359 gaattgctctacgggggagcccccgttggcagaatltgcttgcacataagtgctgat 2418  
Db 5296 GAATTGCTTACGGGGGAGCCCCCGTTGCGAAGAAATGCTTTGACATAGTGTGAT 5355  
OY 2419 gtgtgtcttccagaccttgtctctaggtcctaagcgctgagtggtgagcagca 2478  
Db 5356 GTGCTGCTTCCAGCCCTTGTCTTAAAGCTCAAGCGCTCGAAGTGGGAGGACGCA 5415  
OY 2479 gctgtgagacagctccggggagccttactagaaggaagacttgaattcacaatgctc 2538  
Db 5416 GCGTGAAGAGACGCTCCGGAGGCGCTACTCAGAGAGAGACTTGAAGTTCCACATCCCTTC 5475  
OY 2539 caagtggagatcgaatcattatgtagacgcaacgtgtgcaagaacactcagagactcgtg 2598  
Db 5476 CAAGTTGGAATTCAGCTCTATGTTAGACGCGCACCGTAGAGAACTCCAGACTCGTGG 5535  
OY 2599 aagggaactatctgtaacttctgacacacacacacacacacacacacacacacacac 2658  
Db 5536 AAGGAGCTTATCTCGTACTTTTGAACACACACAGGCTGTGAATTCAGAGAAATCCCC 5595  
OY 2659 acctgtaacatcatccatccacgtttaagcgcgcacccctccgacttcgggttgaagacc 2718  
Db 5596 ACCTGAGTCCATCATCCATCCAGTTAAGCTGTGCGCCACACCTCCGACCTGGGGTGAAG 5655  
OY 2719 gaaaagactgaaaatcccttaagcttgcctccatcagctgtgttcttacttctgtaac 2778  
Db 5656 GAAAAGACTGAGATCCCTTTAAGCTTCCGCTCCATCGCTGCTTACTTACCTAACAT 5715  
OY 2779 aactcccaagtttaagtaaacgcttgtgagcagcccgaaactccataaaccttacc 2838  
Db 5716 AACTCCCGACGCAAGTAAAGCCTTATAGACAGCTGGAACCCCATAGACCTTTATC 5775  
OY 2839 tctcaactgttacttactgactccgttacaaggtatataatataacagcactcaagggga 2898  
Db 5776 CTTTACCTGTGCTGATTATTGATGAGGCTTATGAGGCTGACATGAAATAGACATCGAGTGT 5835  
OY 2899 ggtctcccttggggagctgtgtgcctgaatataatgtcgtccttgcatacgaataccctg 2958  
Db 5836 TGCTCTTGAAGGACCTGTGGGCTGACATGCTTCTCCTCCGATTAATTAACCCCGC 5895  
OY 2959 tctcaatgaccagcagcaccaccccgatgtactcgtgtgcttgaaggtttaaagttggcc 3018  
Db 5896 TGTTA-----AAGACACACCTCCCACTAGTCCGTATGATGGGTTCTATTGCTGCC 5949  
OY 3019 aggaccccaataatgaagaatlatgtggaatccctcaggaatlttcccttgaagaatg 3078  
Db 5950 AGG---CACAGAGAAGAGAAATATCTGTGGGGTCTGTGGGAATCCTTGTGAGAGAG 6006  
OY 3079 gagctgataacttcaatgagtggaatgtggaatgtgcagctcctcagcaagaagag 3138  
Db 6007 GAGCTGCTGACCTCCCAAGATGAGACTGGAATAGGCGGATCTCTCCAGAGACGGGT 6066  
OY 3139 aagttacttctgttaa 3156  
Db 6067 AAATTTCTCTTTGTCAA 6084  
|||||





Db 3810 ATAACTGTAAATAGCCCCCATGATGAGAAACATGTTGGAGAGCCCCCAGACCGATG 3869  
Qy 919 atgaccaacgcccgaatgaccacatacaagcctgctctcaagaagggctcacttc 978  
Db 3870 ATATACCAACGCCCGCATGACCACTATCAAGCCTGCTTCTCACAGAGGGGTCACTTTC 3929  
Qy 979 gctccacacgcccctcacaacctgcaactctctgctcgaagagactaagaacagt 1038  
Db 3930 GCTCCACACGCGCTCTCAACCCCTGCACTCTTCTCCGGAAGAGACTATATACCACTG 3989  
Qy 1039 actcagatctgcataactatltgaltgagagactgggtcgcgaagaccttacaag 1098  
Db 3990 ACTCAGATTTGCATCACTATTTGATTGAGAGACTGGGGTCCGCAAGACCTTACAGAC 4049  
Qy 1099 ataccgctacgaggaagtgctaaactggttactgctcgaagagagactatgtgtgaa 1158  
Db 4050 ATACCCCTACTGAGAAAGTCTTAACCTGTTCTACTGACGGAAGAGCTATGTTGTGAA 4109  
Qy 1159 ggtlaagagatgctgctgagcgagtggtgagacgagccgacagatctgagcgagc 1218  
Db 4110 GGTAAAGAGATGCTGGGGCGGCACTGTTGAGAGGGAGCCGACATCTGGGCGACAGC 4169  
Qy 1219 ctgcccgaaggaacttcagcgcaaaagctgagctcagctcagcagccttctgcyg 1278  
Db 4170 CTGCGGAGAGCACTTCAAGCGCAAAAGCTGAGCTCATGGCCCTCAGCAAGCTTTGCGG 4229  
Qy 1279 ctgcccgaaggaagatcacaataatcattatagagacagatagatgctcttgagcga 1338  
Db 4230 CTGCGCGAAGGAATTCATTAACATTTATACGACAGAGATGAGCTTTGCGCATGCA 4289  
Qy 1339 caagctacacgaggccatcatalaataaagaagggtgtgtactacacagcgaggaata 1398  
Db 4290 CACGTACACGGGGCATCATATTAACAAAGGGGTTGCTTACTCAGCGAGGAGGAATA 4349  
Qy 1399 aagaacaaagggaatcttaagctatlaagaaccttaattgccaagaagctagct 1458  
Db 4350 AAGAAACAAGAGAAATTTCAAGCTTATTAAGAGCTTACATTTGCCAAAAGGCTAGCT 4409  
Qy 1459 attatacactgtcctgagacatcaagaagcaaatcatalatlaaggaggaagagt 1518  
Db 4410 ATTATACACTGCTCGACATCAGAAAGCCAAAGTCTCATATCTAGAGGGGAACCAAGT 4469  
Qy 1519 gctgacgggtgtgccaagcaagcgccagcgtgtaacctctgctcatalatagaagc 1578  
Db 4470 GCTGACCGGTTGCCAAGCAGGAGCCAGGCTGTTAACTTCTGCGCTATATATGAAGC 4529  
Qy 1579 cccaaagcccaagaacccaagacagatcaacctagaagacgtggaagagataaaag 1638  
Db 4530 CCGAAAGCCCGAAGCCAGACGACAGTACACCTAGAAAGACTGGCAAGATTAATAAG 4589  
Qy 1639 atagaccagtctctggaactcggagggagccgtctataacctatataatggagaatc 1698  
Db 4590 ATAGACCGAGTTCTCTGAGACTCCGAGGGGACCTGCTATACCTCATATGGAAGAAATC 4649  
Qy 1699 ctgcccacaagaagggttagaatatgtccaacagatacatgtctcaaccacactaga 1758  
Db 4650 CTGCCCAACAAGAGGGTTAGATATGTCCAAGATAGTGTCTTAACCCACCTTAGGA 4709  
Qy 1759 acttaaacactgcaagcgtgtgtcaagaacatccctatcatgttctgaagcttacaaga 1818  
Db 4710 ACTAAACACCTGACGAGCTGTTGAGAAACATCCCTTATCATGTGTGAGGCTAACGGA 4769  
Qy 1819 gtggtctgactcgtgtgcaaacattgtgtgctcgtcgaagctgttaagtactccctcc 1878  
Db 4770 GTGCTACTACTGGTGTCAACATTTGTGTCCTGCGCAGTGGTTATGTATATCTTCC 4829  
Qy 1879 agaatacctccagaaagagactlaaagggaagccaccagcgctcaactggaagtgagc 1938  
Db 4830 AGAATACCTCCAGGAAGAACTTAAGGGGAAGCCACCCAGGCGCTTACTGGGAAGTGAGC 4889  
Qy 1939 ttcactgaagtaaacgagctaaataatacgaacaaatlatctatgttcttgtagaac 1998  
Db 4890 TTCACTGAGGTAAGCGGCTTAATATAGGAACAATATATGTATTTGTTGTAGACACC 4949

Qy 1999 ttccaagatgggttagaggtcttaccataagaagaagacttcaacctgtgtgtgctaag 2058  
Db 4950 TTTTCAGGATGGGTAGAGGCTTATCTACTTAAGAAAGAGACTTCAACCGTGGGTAG 5009  
Qy 2059 aaatctcggaggaatcttccaaagatttggaaatcactaagtgaataaggtcagaacat 2118  
Db 5010 AAAATACTGGAGAAATTTTTCAGATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAAT 5069  
Qy 2119 ggtccagcttctgtctccagaglaagtcaggaactgccaagatattggggattgtatg 2178  
Db 5070 GGTCCAGCTTCTGTTGCCAGATAGTCAGAGAGCTGGCCAAAGATATTGGGATTTGATTGG 5129  
Qy 2179 aaactgcattgtgcatacagaaccccaagctcaggaaggttagagagatgaataagacc 2238  
Db 5130 AAATGATGTTGATACAGACCCCAAGCTCAGAGACAGGTAGAGAGATGATAGAAC 5189  
Qy 2239 attaaagagcccttaccaaattgacacacagagactggaacttaagtatgtgtctc 2298  
Db 5190 ATTAAAGAGACCTTACCAATTTGACCAAGACAGCTGGCATTTATGATTTGATGGCTCTC 5249  
Qy 2299 ctgccccttctgtcttctttagagtgaggaacacccctggaagcttggctgaccctat 2358  
Db 5250 CTGCCCTTGTGCTTTTAGGCTGAGAGAACCCCTGAGAGTTGGCTGACCCCTAT 5309  
Qy 2359 gaattgctacaggggagaccccccgttggcagaatgtgcttgcacatagtgctgat 2418  
Db 5310 GAATTTGCTTACGGGGAGCCCGGCTGGAGAAATTTCTTTGACATAGTGTGAT 5369  
Qy 2419 gtgctgtcttccagcccttctctctctctctctctctctctctctctctctctctct 2478  
Db 5370 GTGCTGTTTCCAGACCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5429  
Qy 2479 gcttgaagacagctccggagagcctaactcagggagagacttgaagtccacatgcttc 2538  
Db 5430 GCGTGAAGAGAGCTCCGGAGGCTTACTCAGAGAGAGACTGCAATTTCCACATCCCTTC 5489  
Qy 2539 caagtttggagatcagctctatgttgaagcagcgtgcaagaaacctcgaagactcggtg 2598  
Db 5490 CAAGTTGGAGATTCAGTCTATTTAGACGCCACGTCGAGAAACCTCGAGACTCGGTG 5549  
Qy 2599 aagggaacctatctgcttcttctctctctctctctctctctctctctctctctctct 2658  
Db 5550 AAGGACCTTATCTCTGACTTTTGGACACACCAAGCGGTGTAAGTCAAGAAATCCCC 5609  
Qy 2659 acctgatalcatgcatccacgttlaagccgagccaccccgatctgggttgaagacc 2718  
Db 5610 ACCTGGATCCATTCATCCACGTTAAGCGGGGCCACCTCCGACTCGGGGTGGAGAGCC 5669  
Qy 2719 gaaagactgaaatcccccctaagcttgcctccatcgctgtgttctctactgtlcaat 2778  
Db 5670 GAAAGACTGAGAAATCCCTTAAAGCTTCCCTCCATCGCTGCTGTTCTTACTTAACAT 5729  
Qy 2779 aactcctcaagttaagtgaagacgcttggagagcccgaaacctccataaaccttacc 2838  
Db 5730 AACTCCCGAGCGAGTAAGTAACGCTTATTAACAGCTGGAACCCCAATAGACCTTATTC 5789  
Qy 2839 tctcaacctgttactactcagctccgagctcagagatlaataatacaacacactcaaggga 2898  
Db 5790 CCTTACTGCTGATATTATTTAGACCTTGATACGGGTGTCACTGTAATTAACACTCGAGGT 5849  
Qy 2899 ggtcccttggagacgtgtgtgctgaatlatatgtcgtccttgcataagtaaatcccg 2958  
Db 5850 TGTCTCTAGAGGACCTGTGTGCTGAGACTGATTTTTCGCTCGATTTGATTAAACCCGC 5909  
Qy 2959 tctcaatgacagcgcaaccccccgatgtactcgtgttcaaggtttaaogtttgc 3018  
Db 5910 TGTTA-----AAACACACCTCCCAACCTTACCTGATTTATGCTTATGCTGCC 5963  
Qy 3019 aggaccccaataatagaagatattgtggaatccctcaagattccttgcagcaatg 3078  
Db 5964 AGC-----CACAGGAAGAAATTAAGTGTGGGCTTGTGGGAATCTTCTGTAGAGATG 6020

QY 3079 gagctgcataactcttaataatgaggaattggaatgagccagctctcagcaagagagt 3138  
|||||  
Db 6021 GAGCTCGTCACCTCCAGAGAGACTGGAATGCGGATCTCTCCAGAGCCGGGT 6080  
QY 3139 aagtaactctttgttaa 3156  
|||  
Db 6081 AAAATTCTCTTGTCAA 6098

RESULT 9  
ARI30475  
LOCUS ARI30475 8132 bp DNA PAT 16-MAY-2001  
DEFINITION Sequence 3 from patent US 6190861.  
ACCESSION ARI30475  
VERSION ARI30475.1 GI:14118800  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 8132)  
AUTHORS Fishman,J. A.  
TITLE Molecular sequences of swine retroviruses method of using  
JOURNAL Patent: US 6190861-A 3 20-FEB-2001;  
FEATURES  
Location/Qualifiers  
1..8132  
BASE COUNT 2248 a 1977 c 2037 g 1870 t  
ORIGIN

Query Match 87.8%; Score 2915.2; DB 6; Length 8132;  
Best Local Similarity 94.3%; Pred. No. 0;  
Matches 3051; Conservative 0; Mismatches 173; Indels 12; Gaps 2;

QY 22 gatgcctctcttgctcgtgataacacccactaagcaacacttttgctctgaatg 81  
|||||  
Db 2970 GATGCTTCTTCTGCTGTGATTACACCCACTAGCCAACTTTTGCCTTCGAATGG 3029

QY 82 agagatcaggtacggaagaacccggagctcactgagaccgactgccccaaaggttc 141  
|||||  
Db 3030 AGAGATCCAGGTACCGGAAGAACCGGAGCTCACCCTGACCCGACTGCCCCAAGGGTTC 3089

QY 142 aagaactcccgacacactcttgacgaagccctacacagggactgagccaacttcagatc 201  
|||||  
Db 3090 AAGAATCCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGATTC 3149

QY 202 caaacacctcaagtacacctctctcagtaagtgatgagctgctctggcggaagcaac 261  
|||||  
Db 3150 CAACACCTTCAGAGTGAACCTCTCTCCAGTACGTGATGACCTGCTTCTGCGGAGCCAC 3209

QY 262 aacacgactgcttaagaagtgacgaagcactactgctggaattgtctgacctaggtac 321  
|||||  
Db 3210 AAACAGAGATGCTTGAAGAGCACGAAGGACTACTGCTGGAATTTGCTGACCTAGGCTAC 3269

QY 322 aagacctctgtaagaagcccaagattgcaagagagagtaacataacttgggtacagt 381  
|||||  
Db 3270 AGAGCTCTGTCTAAGAGGCGCAGATTGTCAGAGAGAGAGTAACTACTGGGGTACAGT 3329

QY 382 ttggcgaggcgagcagctgagtgctgacgaagagcaggaagaacactgtagtccagtlaccg 441  
|||||  
Db 3330 TTGGCGGAGCGGAGAGGCTGTGACGGAGGACGAAAGAAACTGTGCTCAGATACCG 3389

QY 442 gccccaacacagccaacaaagttagagagtttttgggagacagctggaattttgacagctg 501  
|||||  
Db 3390 GCCCAACACACAGCCAAATAGAGAGTTTTGGGAGACAGCTGCAATTTTGCAGACTG 3449

QY 502 tggatccgggggtttggaaccttaagcccaactctaccogctaaaccaagaagaagg 561  
|||||  
Db 3450 TGGATCCCGGGGTTTGGCAGCTTACAGCCCACTTAACCCCTAACCAAGAAAGAGG 3509

QY 562 ggaattctcgtcctcctgaagcagaagaaggaattgtagtataaagaagccctgctg 621  
|||||  
Db 3510 GAATTCCTCTGGGCTCTGAGCACCAGAGAGGCAATTTGATGATCAAAAAAGCCCTGCTG 3569

QY 622 agcgacactgctctgcccctccctgacgtaactaaaccccttaactttagtggatgag 681  
|||||  
Db 3570 AGCGACCTGCTCTGGCCCTCCCTGACGTAACCTTAACCCCTTTATGTGATGAG 3629

QY 682 cgtaaaggagtagcccgaggaagtttaacccaacccctagaagacatgagagagactgtc 741  
|||||  
Db 3630 CGTAAGGAGTAGCCCGGGGAGTTTAAACCAACCCCTAGGACCATGAGAGAGACTGCTC 3689

QY 742 gcttaactgtaagaagctgacccgtagccagtggttgccggatgtctgaagct 801  
|||||  
Db 3690 GCTTACTCTTAAAGAAAGCTGTGATCTGTACCGAGTGGGCCCATATGCTTAAGGCT 3749

QY 802 atcgagctgtygccaactactgtaacgaagcgttaccaattgacttgggacaagaata 861  
|||||  
Db 3750 ATCCGAGCTGTGGCCATAGTGTGTAAGAGACCCCTACCAATTTGATTTGGGACAAATATA 3809

QY 862 actgtaalagccccccatgcatlgagaacalcgttcggcaagcccccaagccgaatgag 921  
|||||  
Db 3810 ACTGTAATAGCCCCCATGATGATGAGAACATCTGTCGGCAGCCCCAGACCCGATGATG 3869

QY 922 accaagcccgcatgaccacatcaagaagcctgtcttcaagagaggtcacttgcct 981  
|||||  
Db 3870 ACCAAGCCCGCATGACCCACTATCAAGCCTGCTTCAACAGAGAGGTCACGCTGCT 3929

QY 982 ccacacagccgctcacaacctgcccactctctgctgtaagaagatgataccaagtact 1041  
|||||  
Db 3930 CCACACAGCCGCTCTCAACCCCTGACCTCTTCTGCTCAAGAGATGATGAAACAGTACT 3989

QY 1042 catgattgcaatcaactatgattgagagagctggttcgcaagaaccttaagacata 1101  
|||||  
Db 3990 CATGATTGCCATCAATATTGATTGAGAGAGACTGGCTCCGCAAGCACTTACAGACATA 4049

QY 1102 ccgctgactggaagaagtgtaacctggttaactgacgaagagcgtatgtggtggaggt 1161  
|||||  
Db 4050 CCGCTGACTGGAAGAGTCTAAGCTGTGCTCACTGACGGAAGACACTATGTGTGGAGAGT 4109

QY 1162 aagagatgctggtggcgagtagtgtaacggagcccgacagatctggtcgcaagcctg 1221  
|||||  
Db 4110 AAGAGATGCTGTGGGGCGGCGGTGTGAGCGGAGCCGCAAGATCTGGCGACAGGCTG 4169

QY 1222 ccggaaggaacttcaagcgaagaagctgagctcaatggccctacgcaagcttggcgctg 1281  
|||||  
Db 4170 CCGGAAGGAACCTTACGACCAAAAAGCTGAGCTCATGTGCGCTTCAAGCAAGCTTGGCGCTG 4229

QY 1282 gcggaaggaataatcataaacttataagacagcaggtatgcttggagctgcaac 1341  
|||||  
Db 4230 GCCGAAGGAATCATTAACATTATACGACAGCAGGTATGCTTTGGACTGCACAC 4289

QY 1342 gtacacggggcatalaacaagaagggttgccttaccctcaagcaggaaggaataaag 1401  
|||||  
Db 4290 GTACATGGGGCATCTTAACCAAAAAGGGGTGCTTACCTCAGAGGAGGAGAAATTAAG 4349

QY 1402 aacaagaaggaattctaaagcctatagaagccttaccattgccaagaagctagctatt 1461  
|||||  
Db 4350 AACAAAGAGGAATTTTAAGCCTATTAGAGCCCTTACTTATTAACCAAAAAGGCTATTT 4409

QY 1462 ataacactgctgacatcgaagaagccaagaatctatctcgaaggaagcaagatgct 1521  
|||||  
Db 4410 ATACACTGCTCTGACATCAAGAAAGCTTAAGATCTCATATCAAGAGAAACCGAATGCT 4469

QY 1522 gaccgggttgcgaagcagcagcccgctgttaaccttgcctataatagaaacccc 1581  
|||||  
Db 4470 GACGGGGTGTGCAAGAGGAGGCCCAAGGTGTTAACCTTCTGCTTAATATAGAAATGCC 4529

QY 1582 aaagccccaagaaccagaagcagtlacacccctagaagaactggaagaagataaagaagata 1641  
|||||  
Db 4530 AAAGCCCAAGAACCAAGACAGCACTACACCTTGAAMAACTGGCAAGAGATTAAGATA 4589

QY 1642 gaccagttcttgagactccgagagagactgctatccctacataggaagaagaaatccgtg 1701  
|||||  
Db 4590 GACCAGTTCTGTGACTCCGGAAGGAGCCTGCTATACCTCAGATGGGAAGAAATCTCG 4649

QY 1702 ccccaagaagaaggttgaataatgtccaaagatacatcgtctaaacccactaggaact 1761  
 |||||  
 Db 4650 CCCCAAGAAGAAGGTTGAATATGTCACAGATACATCTCAACCCACTAGGAACCT 4709  
 QY 1762 aaacaccctgcagcttgctgaacacatcccttactaattcttgagcttccaggagt 1821  
 |||||  
 Db 4710 AAACACCTGCACAGATTGGTCAGACATCCCTTATCATAGTCTGAGGCTACCCAGGAGTG 4769  
 QY 1822 gctgaactcgttgctgaacacatgtgtcccttgccagcttgctgaacttccatccaga 1881  
 |||||  
 Db 4770 GCTGACTCGTGTGTCAAACATGTGTGCTGCCAGCTGCTGTAATGCTAATCTCTCCAGA 4829  
 QY 1882 ataactccaagaagaagacttaagggaagcaaccagagcgtcacttggaagtgaactc 1941  
 |||||  
 Db 4830 ATGCTCCAGAGGAGAGAGACTTAAGGGAGCCACCCAGCGCTCATCTGGAGAGTGACTTTC 4889  
 QY 1942 actgagtgtaaacgagcgaactaatacgaacaataatactattggttttggtagaacctt 2001  
 |||||  
 Db 4890 ACTGAGGTAAAGCCGGCTTAATACGGAACAAATACCTATTGTTTGTGTAGACACCTTT 4949  
 QY 2002 tcaagatgtgtagaggtcttactactactaagaagaagacttcaacgctgtgtcgaaga 2061  
 |||||  
 Db 4950 TCAGGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGTGGTGAATAAAA 5009  
 QY 2062 atactcgaggaagaatttccaaagatttggaatacctaagtaataaggtcagaagaatggt 2121  
 |||||  
 Db 5010 ATACTGAGAGAAATTTTTCAGAGATTGGAATACCTAAGGTAAATAGGCTCAGACAAATGCT 5069  
 QY 2122 ccaagcttcgttgcccaagtagaagtagagacttgcccaagatttggtgagatttgtaa 2181  
 |||||  
 Db 5070 CCAGCTTTTGTGTCAGGATGAGTCAAGGAGCTGGCCAAAGATATGGGATTTGATTGGAAA 5129  
 QY 2182 ctgcattgtgatacagaaccccaagctcaggaacaggttagaagaatgaataaacat 2241  
 |||||  
 Db 5130 CTGCATGTGTGATACAGACCCCAAGCTCAGAGACAGTAGAGAGATGAATAGAACCAT 5189  
 QY 2242 aaagagacccttaccacaatttgcacacagagactgagacttaattgattgagctctctg 2301  
 |||||  
 Db 5190 AAAGGAGCCCTTAAATTTGACCGCGAGACTGGCCCTTAAATGATTTGATAGCTCTCTCG 5249  
 QY 2302 ccccttgctgcttttaaggtagaagaacaccccttgaaacagtttgagcagccctatgaa 2361  
 |||||  
 Db 5250 CCGTTTGTGCTTTTAAAGGTTAAGGAACACCCCTTGACAGTTTGAGGCTGACCCCTTAAGAA 5309  
 QY 2362 ttgctctacagggagaccccccgcgttgagcaagaattgcttctgacataagtgctcagtg 2421  
 |||||  
 Db 5310 TTACTCTACGGGGAGCCCGCCCAATTTGTGTAGAAATTTGCTGTACATGATGCTGACGTG 5369  
 QY 2422 ctgctttcccaagccttctgtctcctagctcagagcgtcagagtggtgtagcagagagcg 2481  
 |||||  
 Db 5370 CTGCTTTCACAGCCCTTGTGTTCTTAAGGCTCAAGGCACTTGAAGTGGGTGAGACAAAGAGCG 5429  
 QY 2482 tggagagacagctccggagagcctactc--aagagagagacttgcaagtccaactcgtctc 2538  
 |||||  
 Db 5430 TGGAGGCAACTCCGGAGGAGCTACTCAGAGAGAGAGAGACTTGCGATCCACATCGTTTC 5489  
 QY 2539 caagtgtgagaattcaagtatgtttagaagcagcagcgttgcaagaacccctcgaactcgttg 2598  
 |||||  
 Db 5490 CAAGTGGAGATTCAGTCTACTAGCTTAGAGAGCCAGCGTCAGAGAAACCTCGAGACTCGGTG 5549  
 QY 2599 aagagacttactcgtcgaacttttgcacacacacagagcgttgaaagtgcagaagaatccc 2658  
 |||||  
 Db 5550 AAGGCGCCCTTAATCTGCTACTTTTGAACACCAACAGGCTGTGAAGAGTGAAGATCTCC 5609  
 QY 2659 accgtgattcactgcatcccaagcttaagcagcagcagcactccgaattgggggttgaagcc 2718  
 |||||  
 Db 5610 ACCTGGATCCATGCTCCACGTTAAACGGGCGACCTCCGATTCGGGGTGGAAAGCC 5669  
 QY 2719 gaaaagactgaataatcccttaagcgttcgctccatcgctgtgttcttactcgtcaat 2778  
 |||||  
 Db 5670 GAAAGAGCTGAATAATCCCTTAAGCTTCGCGCTCATCGGCTGCTTACTACTGCTGAAT 5729  
 QY 2779 aactccccaagttaagttaaaagccttgtagacagccgaactccataaaccttatac 2838

Db 5730 AACCTCAGACCTAATATGATGCGATAGAGAGACACCTGAACCTCCCAATAAACCTTATC 5789  
 QY 2839 tctcaactgttacttacttacttacttacttacttacttacttacttacttacttacttact 2898  
 |||||  
 Db 5790 TCTCACTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5849  
 QY 2899 gctcccttgaggagacgttgagcgtgaatataatgtctgctcctcgaatcgaatccctg 2958  
 |||||  
 Db 5850 GGGCTCTTAAAGAACTCGTGGGCTATCATACGTTTGGCTCAAGTATTAATTTCTGAG 5909  
 QY 2959 tctcaatgaccagagccacaccccccgatgactcgttgcttactgaggttcaagtttgcce 3018  
 |||||  
 Db 5910 TCT-----GACCTCACCCCGAGATATCTCCATATGCTCAAGGATTTATTTGTCGCC 5960  
 QY 3019 aggaaccccaataataaagaatatttggtaaatccctcagagtttcttcttgcagaagatg 3078  
 |||||  
 Db 5961 AGGACCACTAATAATTAAGAAACATTTGCGGAATATCCAGAGATTTCTTTGTAACAAATG 6020  
 QY 3079 gactgcataacttctaatgtatgtggaatttgaatgtgagccagctcctcgaagaagagat 3138  
 |||||  
 Db 6021 GAACGTGTAACTCTAATATGATGATTTGAATATGCGCAACCTCTCACAGAGATAGGT 6080  
 QY 3139 aagttactcttggtaacaaatccatccaggtataatcaatlaattatgagcattgagag 3198  
 |||||  
 Db 6081 AAGTTTTCTTATGTAACACACTATACAGCTCTGCAATTAATTAATTAATTAATTAATTAAT 6140  
 QY 3199 atggaagatttgcaacagcgggttacaagaagtgtgacgaatlaagcaataagct 3254  
 |||||  
 Db 6141 TAGAATCTGAAGGCCCAAGTGTCTCTTCAAGACCTAATTAATTAATTAATTAATTAAT 6196

RESULT 10  
 LOCUS AF038600  
 DEFINITION Sus scrofa porcine endogenous retrovirus PERV-MSL mRNA, complete sequence.  
 ACCESSION AF038600  
 VERSION AF038600.1 GI:3133301  
 KEYWORDS  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 8132)  
 AUTHORS Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.  
 TITLE Identification of a full-length cDNA for an endogenous retrovirus of miniature swine  
 JOURNAL J. Virol. 72 (5), 4503-4507 (1998)  
 MEDLINE 98216827  
 REFERENCE 2 (bases 1 to 8132)  
 AUTHORS Fishman,J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REFERENCE 3 (bases 1 to 8132)  
 AUTHORS Fishman,J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REMARK Sequence update by submitter  
 COMMENT On May 16, 1998 this sequence version replaced gi:3116441.  
 FEATURES  
 source  
 1..8132  
 /organism="Sus scrofa"  
 /strain="miniature swine"  
 /db\_xref="taxon:9823"  
 /cell\_type="lymphocytes"  
 /note="porcine endogenous retrovirus PERV-MSL"  
 LTR 1..564  
 CDS 585..2159  
 /codon\_start=1

/product="gag protein"  
/protein\_id="AAC16763.1"  
/db\_xref="GI:3116442"  
/translation="MGQITPTPLSLTLDHMTPEYKSAHNLSTVQYKKGPMOTCYSEMP  
TPDYGNPSESTSEITLAKAVITQDGHSDDEPTLTITWDLDEDPWPWKLWLN  
KPKPGRIILAEENKHSKVEKPSPHLYPELEPPAPMEDESVPPYLVQAGAR  
FPAPGAPAVEGPAAGTNRKRGATPERTDEIATLPTGTGPPPGGOLQPLYPFSS  
ADLYNMKTNHPESEDPORLTGLVESLMSHOPTWDCOOLQTLTTEERERILIEA  
RKAVPADGPRITLONEDIMGFPRLTGMDYNTAEERESLKITIROLAVLITGSAIRP  
TNLAKRYEMOGPNBPSPVLEPRLIAPRPPDPTNPKRASAVALITGSAIDIR  
KTIQRLGLOEALRDLVYAEKVIYTKRETEERERDRKREBERERENKQEKVL  
KILAAVEGKSNTEREDPRKTRSGPRSGNLGNRTPLDQCAVCYCKERGHWARNCPK  
KGNKGRILALEEDND"  
2307..5744  
/codon\_start=1  
/product="pol protein"  
/protein\_id="AAC16764.1"  
/db\_xref="GI:3133302"  
/translation="MGATGQOQYPMWTRRTVDLGVGRVTHSFLVPECPAPILGRDL  
TKMGAQISEQCKREVSANNKPTVTLQDDEYRLYSPLVBDONIQWLEPQDAM  
AETAGMGLAKOVPVOYLKASATPVSVROYPLSKAOGIRPHVORLQOGLILVPO  
SPMNTPLPRKRGKNDYRPVODLREYNKRVODIHPTVNPNTILCALPQSSWTVL  
DLKDAFECRLHPTSPPLAFEMRPDGTGRTGQLITRLPQCFKNSPTIFDALRDL  
ANFRIOPOVTLQIVYDDLLAGATKQDCEGKALLLESDLYRASAKKQICRE  
VYTLGLSLNDGQWMLTEARKKTVQVLPAPTAKOMERLGTAGFCRLMIPGATLAP  
LYPLTEKEGEFSWAPRHOAFDAIKALLSAPALADVTGKPTLYVDERKIVARGVL  
TOTLGPWRBPVALSKLDPVASGMPICLKALIAVAALIVKADKLITLGNITVAPHA  
LENTVROPDRMNTNARMTHYOSLITLERTFAPAPALINAPDLIPREDEPTHCHO  
LLEETGVKRLDLPITGEVLTWPTDSSYYVEGRNAGAAVVDGTRITMASLPGG  
TSAOKAELMALTOALRLAGKSNITVTSRVAFAHVAIGALYKQGLTTSAGRIK  
KEBILSLLEAVHLPKRLAIHCGHQRADLISRGQMDRVAKAQAOVNLPTTEM  
PKAPERRQVLTLEMDQEIKKIDOFSETPGTCYSDGKELILPKEGLEVVQO IHRLLH  
LGTKHLQOQVARTSPYHRLPGVADSVNHCYCOLVNPSPMPGKRLRSGHGAH  
MEYDFTEVPKAGKNYLLVFDYDSGWEAVPTKSTSTVAKKLLEPIPRFPIK  
VIGSDNGPAPVAHVSOGLAKIIGIMKILCAARPOSSGVEEMNR IKETILKILAFET  
GVNDWIALPLPVLFRYRNPQCGFLPTLYLIGGPELIVASHADVLQSPLESR  
LKALEWNRQRAMSOLKEATSGGDDQIPIRFQVGSVYVRRRAHGLTRKMGAPYLV  
LPTPTAVKVGISTWIMASHVKKPAPPDGMAEKETENPLKLRHVVYVSNNLSD"  
5620..7536  
/codon\_start=1  
/product="env protein"  
/protein\_id="AAC16765.1"  
/db\_xref="GI:3133303"  
/translation="MHPILNRHLPIRGKPKRLKIPLSFASIAFWLTLSTQNGM  
RIEDSLNSHKPLSLTLITDSTGINNTQCEAPITGMYPLVLCRLDYISLSP  
DILHAGFYPCGPNNGKHGNDKPFCKOMKNTGNSNGYKPLQSDRRESEFVN  
TYSGQFNYLIMIRGSPKSPSDDLYKISTEKGQENLILKYNKSMWAMYVGG  
SKQPSILITRIKINOLEPMAIGPNYLTQRPQGPSSNTISGDSPTESST  
TKMGALFSLIQGAFOLNSTPEATSCWLCIASPPYEESMARGENVYKERRDQ  
CTWGSQNKLTLENOTKPCIMVOIVRYVYPEKALIDEDYDRNHRKREPLTLAV  
MGLTGVASGDTGTALVTPQOLETGLSLNHRITEDOLNMLEKSVNSLESTLSE  
VVLONRGGDLLEFKEGGLCVALKKECCGYVDSGASIRSNMALKLEPLRREKET  
OGMEEGMNRSLMALTLALSALTGPLVLLLTITVPCITIKLIAITRERISAVQIMVL  
RQOYQSPSSREMR"

LTR  
BASE COUNT 2247 a 1977 c 2037 g 1871 t  
ORIGIN

Query Match 87.7%; Score 2912; DB 4; Length 8132;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 3049; Conservative 0; Mismatches 175; Indels 12; Gaps 2;

QY 22 gatgcctcttcgctgaattacacccactagcaacacttttccttcgaatg 81  
Db 2970 GATGCTTCTTCTGCTGGAATTACACCCACTAGCAACACTTTTGTGCTTCGATGG 3029  
QY 82 agagatccaggtacggaagacccgagctcaactggaaccgactgcccagaagttc 141  
Db 3030 AGAGATCCAGGTACGGGAGAACCCGGGACACTCAGCGACCCGACTGCCCAAGGTTTC 3089  
QY 142 aagaactccccgaacatctttgaagaagccctacagaaggaactgccaacttcagatc 201

Db 3090 AAGAACTCCCGGACCATCTTTGACGAAGCCCTACAGAGAGACTGGCCAACTTACGAGATC 3149  
QY 202 caaacccctcgaagtgcacctctccagtaagtgtatgacctgcttcctggcggaagccacc 261  
Db 3150 CAACACCCCTCAGGTGACCTCTCCAGTACGGGATGACCTCTTGTGGCGGACCAACC 3209  
QY 262 aaacagaactgtctaaagaatcgaagcaccactgtcgaattctctcactaagctac 321  
Db 3210 AATCAGACTGCTTACAGAGCAGAGAGCACTACTGCTGGAATTTGTTGACTAGGCTAC 3269  
QY 322 agagcctctgctaaagaagcccaagatttcgaagagagaggttaactatctggggtacagt 381  
Db 3270 AGAGCCTCTGCTAAGAGGCCCAAGATTTCAGAGAGAGAGCTTAACATCTGGGGTACAGT 3329  
QY 382 ttgcggagcgagcgagcgtatgctgcagagagcagcaggaagaaactgtagtcagatccg 441  
Db 3330 TTACGGGAGCGGCGAGGATGCGTGCAGCGAGCGGAAACGATGATCCAGATACCG 3389  
QY 442 gcccccaaccacagcaacaagtgagagagttttggggagagcgtggaatttcgaagctg 501  
Db 3390 GCCCCACACACAGCCAAACAAATGACAGAGTTTTTGGGACAGCTGATTTTGCAGACTG 3449  
QY 502 tggatccgagggtttgagaccttaagcagcccaacttaaccgctlaaccagaagaaaggg 561  
Db 3450 TGGATCCCGGGGTTTGCGACTTAGAGCCCACTTACCCGCTAACCAAGAAAGAAAGG 3509  
QY 562 ggaattcctctgagctctcgtgagcacaagaagagcatttgatgctatacaaaaagccctgt 621  
Db 3510 GAATTCCTCTGGGCGCTCCAGCAGCAAGAGGATTTGATGATCAAAAAGCCCTGCTG 3569  
QY 622 agcgacccgtgccttgagccctccctcgtacgttaactaaccttacccttatgtgagtag 681  
Db 3570 AAGCGACCTGCTGCTGGCCCTCCCTGACGTAACCTTAACCTTTATGTGATGAG 3629  
QY 682 cgtlaagagagtagcccgagagagttttaaaccacccctlaagacatlgagagagacctgt 741  
Db 3630 CGTAAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGACATGAGAGAGACTGTC 3689  
QY 742 gctctcgtctcaagaagagttgactcgtgaagcaatggttgccgtatgtctgaagct 801  
Db 3690 GCTTCCTGTCGCAAGAGCTTCATCTGTAGGCAATGTGTGGCCATATGCTTGAAGGCT 3749  
QY 802 atcgagcgtgtgcatctatggttcaagaagcgtgtaacatcttggagacagaata 861  
Db 3750 ATCGCAGCTGTGCGCATATGCTCAAGAGAGCTGACAAATTGACTTTGGGACAGATATA 3809  
QY 862 actgttaatagcccccatatgcatlttgagaacaacatcgcttcgagccccagacagatgag 921  
Db 3810 ACTGTAAATAGCCCCCATGATGAGAAACATCGTTGCGGACGCCCGACGAGATGATG 3869  
QY 922 accaagcgcggaataccacactatcaagagcctgcttcacagagaaggttaacttgcct 981  
Db 3870 ACCAAGCGCCGATGACCCACTATCAAAAGCCGTCTTCACAGAGAGGCTGACGTTCCCT 3929  
QY 982 ccaccagccgctcacaacctgcacactcttcgctcgaagagacgtgaacacagtgact 1041  
Db 3930 CCACGACCGGCTCTCAACCTCTCCACTCTTGCCTGAAGACATGATGAAACAGTACT 3989  
QY 1042 catgattgcatcaactatgtatgaagagagctgggttcgcaagagacttaccagacata 1101  
Db 3990 CATGATTGCAATCAACTATTGATTGAGAGAGACTGGGGTCCGGAAGACCTTAACAGACATA 4049  
QY 1102 ccgctgactggaagagtgctaacctggttcacttgacggaagacagctatggttgaaagt 1161  
Db 4050 CCGCTGACTGGAGAGTGTAACTGTGTTCACTGACGAGACGATGATGTGTGGAAGT 4109  
QY 1162 aagagatgagcttgaggcgagtggtgagcagagaccgcaagatctggcgcaagcctg 1221  
Db 4110 AAGAGATGAGCTTGAGGCGGCGGTGTGTGACGAGGAGCCGACGATTTGGGCCAGACCTTG 4169  
QY 1222 ccggaaggaacttcagcgcaaaaggtgagctcatgagccctcgaagagcttgcgagctg 1281

Db 4170 CCGGAGAGAACTTACGACAAAGGCTGAGCTCAWGGCCCTCAGCAAGCTTTGGGCGTG 4229  
 Qy 1282 gccgaaggaatccataaacaattatatacagacagcagtaagcttgcagctgcac 1341  
 Db 4230 GCCGAGGGAATCCATTAACATTTATACGACAGAGATATGCTTTGGAGCTGACAC 4289  
 Qy 1342 gtacacgggacatctataaacaagggtgtcttaacctcgaaggaaggaaataag 1401  
 Db 4290 GTACATGGGGCATCTATMAACAAAGGGGTTGCTTACCTACGACGAGGAGAAATGAAG 4349  
 Qy 1402 aacaagaagaatcttaagcctatagaaagccttacatcttgcgaagaaggctacatt 1461  
 Db 4350 AACAAAGAGAAATTTCTAAGCCTATTAGAACCTGACATTTACCAAAAGGCTAGCTATT 4409  
 Qy 1462 atactatgtccttgacatacagaagccaagatcatalctagaagggaacagatgct 1521  
 Db 4410 ATACACTGTCCTTGACATCAGAAAGCTAAGATCTCATATCCAGAGAAACCAAGATGCT 4469  
 Qy 1522 gaccgggttgccaagcagcagccagctgttaaccttgcctataatagaagcgc 1581  
 Db 4470 GACCGGGTTGGCAAGCAGGCGCCAGGCTTAACTTGCCTATATATAGAAATGCCCC 4529  
 Qy 1582 aaagccccaagaccagacagacagtacacctagaagaagctggcaagagataaagaata 1641  
 Db 4530 AAAAGCCGAGAACCCAGACACAGCACTACACCTTAGAAGACTGGCAAGAGATMAAAGATA 4589  
 Qy 1642 gaccagttctctgaagctccgaaggagcctgtatatacctcatataggaaagaatcctg 1701  
 Db 4590 GACCAGTTCTCTGAGACTCCGGAAGGAGACCTGATACCTACATGAGGAAGAAATCTG 4649  
 Qy 1702 ccccaagaagaaggttagaataatgtccaacagatacactgtcttaaccactagaagct 1761  
 Db 4650 CCCCACAAAGAGGGTTAGATATGTCCAAAGATACATGCTTAAACCCACCTAAGAGAACT 4709  
 Qy 1762 aaacacctgcagcagttgttcaagaacatccctatcatgttctgaagctcaagaagtg 1821  
 Db 4710 AAAACACTGCAGAGCTGTGTGACAGACATCCCTTATCATTTCTGAGGCTACCAAGAGTG 4769  
 Qy 1822 gctgaactcgtgttcaaacatltgttgccctgcagctgttgaatgttaactcctcaga 1881  
 Db 4770 GCTGACTCGGTGTGCAAACTTGTGTCTGCTGACAGCTGCTAATGCTAATCCCTTCAGA 4829  
 Qy 1882 atactctcagaagaagacataaggaaggaagccagcagcctacgtggaagtgagcttc 1941  
 Db 4830 ATGCTCTCAGGGAAGAACTAAGGGAAGCACCAGGCGCTCAGCTGGGAAGTGACCTTC 4889  
 Qy 1942 actgaagtaaaagccggtlaaataacgaagaacaaataltcttctgtttttagaacacctt 2001  
 Db 4890 ACTGAGGTAAAGCCGGCTAATACGGAACAATACCTATTGGTTTTGTAGAACACTTT 4949  
 Qy 2002 tcaggaatgggttagagctctactactaagaagaagactcaaccgtgtgtgcctaaagaa 2061  
 Db 4950 TCAGGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGGTGGGTAAAAAA 5009  
 Qy 2062 atactgaagaaatlttccaagaattggaatacctaagtaagtaaggttagaagcaatgtg 2121  
 Db 5010 ATACTGGAAGAAATTTTCCAAAGATTGGAAATACCTTAGGTAATAGGCTAGCAATGCT 5069  
 Qy 2122 ccagcttgcgtcccaagtaagtcaggaacgtgccaagaatltyggatltgaattgaa 2181  
 Db 5070 CCAGCTTTGTTGGCCAGTAAGTCAAGGACCTGGCCAAATATTGGGATTTGATTGAAA 5129  
 Qy 2182 ctgcatgtgtcatagacaccccaagagtcagacaggttagaaggtatgaatagaacatt 2241  
 Db 5130 CTGATTTGTCTATACACCCCAAGCTCAGAGAGGTAGAGGATGAATTAACCAATT 5189  
 Qy 2242 aaagaagcccttaaccaattgacacagagactgacatlaattgattgagctgcctg 2301  
 Db 5190 AAAGAGACCTTACTAATATTGACCGGAGACTGCGCTTATATATTGGATAGCTTCTCG 5249  
 Qy 2302 cccttgtgtctttagagtgagaaacacccctggaagcttggctggaacccctatga 2361  
 Db 5250 CCTTTGTGCTTTTAGGTTAGAGAACCCCTGGAACAGTTGGGCTGACCCCTATGAA 5309

Qy 2362 ttgcttaaggggagcccccgttgcagagaatltgcttgcataagtgctgctg 2421  
 Db 5310 TTACTCTACGGGGAGACCCCCCATTTGTAGAAATTTGCTTGTATATGCTCATGTG 5369  
 Qy 2422 ctgcttcccaagccttgttctcttaagctcaaggcctcagtggttgagcagcgag 2481  
 Db 5370 CTGCTTTCCAGGCTTTGTTCTCTAGGCTCAAGGCACTTGAATGGGTGAGCAACAGCG 5429  
 Qy 2482 tggaaagcagctccgggaagcctacac-aggagaagactltgcaatccacatgcctc 2538  
 Db 5430 TGGAGCAACTCCGGAGGCTTACTAGAGAGAGAGAGACTTCCAGATCCCATGCTTTC 5489  
 Qy 2539 caagtgtgaatcagctatgtttagacgcacacgttgcagaagaacctcagactcgtg 2598  
 Db 5490 CAAGTGGAGATTCAGTACTACGTTAGACGCGCACCGTGAGAAACCTGAGACTCGTG 5549  
 Qy 2599 aaggaagcttactcgttacttltgacacacacaaagcgtgtgaaagctgaaagaaatccc 2658  
 Db 5550 AAGGGCCCTTATCTGCTACTTTTGAACACACACAGCGGCTGTGAAAGTGAAAGATCTCC 5609  
 Qy 2659 acctgatccatgcatcccaagcttaagccggcgccacccctccgacttgcgggtggaagcc 2718  
 Db 5610 ACTTGATTCATGCACTCCACGCTTAACCGCGGCCACCTCCGATTCGGGGTGGAAGCC 5669  
 Qy 2719 gaaaagactgaaatcccccctaagcttgcctcatalcggtgtgttcttacttgcact 2778  
 Db 5670 GAAAGACTGAAATCCCTTAAGCTTGCGCTCCATCGGCTGCTTACTCTGTCAAT 5729  
 Qy 2779 aactcctaagtaatgtgaaagcgttgtgaaagcccgaaatccataaaccttalc 2838  
 Db 5730 AACCTTCTGACTAATAGTATGATGATGATGAGAGAGCTGAACTCCATTAACCCCTTATTC 5789  
 Qy 2839 tctcaactgttacttacttactcagctcggtacagatlaataatlaacagcaactcaagggga 2898  
 Db 5790 TCTCACTGTTAATTACTGACTCCGGACAGAGTATTATATACAAACACTCAAGGGA 5849  
 Qy 2899 ggtcccttgggagcctgtgtgcctgaattalatagltcgtccttgcataagtaacccctg 2958  
 Db 5850 GGTCTCTTATGAGAACCTGTGTGCTGATCATACCTTTGCTGCAAGATGATTTCTCTAG 5909  
 Qy 2959 tctcaatgacagggcgcaagcccccagatgactcgttcttaaggggttttgccttgc 3018  
 Db 5910 TCT-----GACCTACCCCGAGATATCTCATGCTCAGGATTTTATGTTGCTCC 5960  
 Qy 3019 aggaaccccaataatgaagaataltgttgaagaatccctaagatltccttgcagaacatg 3078  
 Db 5961 AGGACACCAATTAATGGAATAATTTGGGAATCCAGAGATTTCTTTTGAACCAATG 6020  
 Qy 3079 gactgtcataacttctaataatgtggaattggaatgtgcccagctcctcagaagaagact 3138  
 Db 6021 GAACGTGTATACCTTAATGATGATGATGGAATGAGGCAACCTTCAGCAGGATAGGCT 6080  
 Qy 3139 aagttactccttgttgaataatcctcaagatlaataatlaattatgtgcagtgag 3198  
 Db 6081 AAGTTTCTTATGTCACACCTATACCAAGCTCTGGAACAATTTAATTTACTGACTGTGAT 6140  
 Qy 3199 atggaagaattggaacagcggtgtacaaaagaatgtacgaataaagcaataagct 3254  
 Db 6141 TAGAACTGGAAACCCCAAGTGTCTCTTCAAGACCTAGATTAACCTAAATAAATGTT 6196

RESULT 11  
 AF038599 8132 bp mRNA MM 21-MAY-1998  
 LOCUS AF038599  
 DEFINITION Sus scrofa porcine endogenous retrovirus Tsukuba-1 mRNA, complete  
 sequence.  
 ACCESSION AF038599  
 VERSION AF038599.1 GI:3133300  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



D	3996	TCATGATGGCATCACTATATGATTGAGGACGTCGGGGTCCGCAAGGACCTTACAGACAT	4055
Q	1101	accgcgtactcgagaaagtgtctaactcgtgttcactctacggaagcagctatgtgttgaaag	1150
D	4056	ACCCCTTACTAGGGAAGTGGTAAACCTGGTTTCACTACGGAACACACTTATGTGTGGAAGG	4115
Q	1161	taagaagatgctcgaggcgagctgtgttgaaacggaaccgcagactgtggccaagacct	1220
D	4116	TAAAGAGATGGCTGGGCGGGGTGGTGGACGGGACCCGACGATCTGGGCCAGAGCT	4175
Q	1281	ggccgaaggaatcotaaaccttatataggaacgacagatgtgctttgagctggacga	1340
D	4236	GCGCGAAGGAAATTCATTAACCTTTATACGGACACAGTAGTAGCTTTGGACTGCACA	4295
Q	1341	cgtaacgaggccatctataaacaagaagggtgtgtcttaacttaacgaaggaggaataaa	1400
D	4296	CGTACATGGGGCCATCTATTAACCAAGGGGGTTGGTTACTCTCAGCAGGAGGAAATTA	4355
Q	1401	gaacaaagagaaatctctaagcctatagaagcctacacttggccaaaaggctagctat	1460
D	4356	GAACAAAGAGAAATCTTAAGCCTATTAAAGCCGTACATTTTCCAAAAAGGCTAGCTAT	4415
Q	1461	tatcacctgctccgtggacatcaagaacgcgaagctcatctatctagaaggaacagatg	1520
D	4416	TATTAATGTGCTGGACATAGAAACCTTAAGATCTCATATCCAGAGAAACGATATGC	4475
Q	1521	tgaccgggttgcgaagcagcagccagcgtgttaaoccttgcctataatagaagcgc	1580
D	4476	TGACCGGGTGGCCAGAGGACCCACGGGTGTTAACCTTTCGCTATTAATAGAAATGCC	4535
Q	1581	caaaagcccaagaccgaagcagcagcagcagccttagaagcctggaagagataaaaaag	1640
D	4536	CAAAAGCCCAAGACCCAGACGACAGTATACACTTAGAAGACTGGCAAGATATTAAGAT	4595
Q	1641	agaccagttctcttgaagctccggaagggaacctgtatatactcataatggaaagaacct	1700
D	4596	AGACCA-TTCTCTGAGACTCCGGGAAGGAGCTGCTATTAACCTCAGATGGGAAGAAATCT	4654
Q	1701	ggcccaaaagaagggttaaatatgtccaac-agaataatgttaaccacacattagaa	1759
D	4655	GCCCCAAGAAAGAGGTTTAATATATCTCCACACAGTACATGCTTAACCCACTAGAA	4714
Q	1760	ctaaacacctgcagcagcttgttgaagacaatccacctatcatgttcttgagctaccagag	1819
D	4715	CTAAMACACTGCAGCACTGGTGCAGAAACATCCCTTATCATGTTCGAGGCTACCAGAG	4774
Q	1820	tggtctgactcgtgtgttcaacatltgtgtgcctgtgcagcttgtttaatgttaactctcca	1879
D	4775	TGGGTGACTGGGTGCTAAACATTTGTGTGCTCCGACGTGGTATATGCTTAATCTCTTCA	4834
Q	1880	gaatacctccaagaaagaagctlaaggggaagccaccgaagcgtcaacttggaaagtggat	1939
D	4835	GAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCAGCGCTCACTGGGAAGTAGACT	4894
Q	1940	tcactgaggtaaagccggtctaaatcacggaacaacatactatgtgttttggtaacct	1999
D	4895	TCATGTAGGTAAAGCCCGCTAAATATGGAACAAATACCTATTGGTTTGTGACACCT	4954
Q	2000	tttaagaatggttagaggtctatccactaaagaagaagcttcaaccttggtgtgtaaga	2059
D	4955	TTTATGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGGTGAGCTTAA	5014
Q	2060	aaatactcggagaaatltttccaagaatttggaaataactaaggtaaataggttaagaaca	2119
D	5015	AAATACTGGAAGAAATTTTTCCAAGATTGGAAATACCTAAGGTAAATAGGCTCAGCAATG	5074
Q	2120	gtccagcttcctgtgccccaggtatgtcagggactggtgccaagatatgtggatgtatgga	2179
D	5075	GTCCAGCTTTTGTGGCCAGTAAAGTCAGGAGCTGGCCAAAGATTATGGGATGATTTTGA	5134

QY	2180	aatgcgatttgcacataagcccccagaagctcagaagacaggtaagaagatgaaatgaacaa	2233
Db	5135	AACTCGATTGTGCTATACAGACCCCAAAAGCTCAGGACAGAGTAAAGAGATGATTAACACA	5194
QY	2240	ttaaagagacccttaccacaaattgacacagagactgcatlaatgatlygaatgctctcc	2299
Db	5195	TTAAAGAGACCCCTTACTAAATTTGACCGGAGAGCTGGCTTATGATGATGATAGCTTC	5254
QY	2300	tgcctcttgctctttaaaggtgtagaagaaccccttgaaagtcttggtgcgcgaccctctg	2354
Db	5255	TGCCCCTTGTCTTTTAAAGGTTAGGAACAACCCCTGGACAGTTTGAGGCTGACCCCTATG	5314
QY	2360	aattgccttaccgagagacccccccgcttgcagaaatitgccttgcacatagtctctga	2419
Db	5315	AATTACTCTACGGGGAGACCCCCCAATGTTGATAAATTCCTCTGTACATAGTCTGATG	5374
QY	2420	tgtctgtctccagccttgctctctcctcagctcgaaggcgtcgaigtggtgtagcagcgag	2479
Db	5375	TGCTGCTTTCCACACCTTTTCTCTTAGGCTCAAGGACCTTAGTGAGGGGAGAACCAACAG	5434
QY	2480	cgtagaagcaactcccgagagccctacac--agagagagacttgcaagttccacatcgct	2533
Db	5435	CGTAGAGCAACTCCGGGAGGCGCTACTCAGAGAGAGAGAACTTGCAGATCCCAKATCGTT	5494
QY	2537	tccaaagttagagatcagctctatgtatgtagaagcagccgltgcagaaacctcgagactcgt	2596
Db	5495	TCCAAGTGGAGATTCAAGTCTACGTTAGAGCCACCGTACAGAAACCTCGAGACTCGGT	5554
QY	2597	ggaagagcaactatctcgtactctttagacacacaaagcgcttgaaagtgcgaagaatcc	2654
Db	5555	GGAAGGCGCTTATCTGTACTTGTTCGACACCAACAGGCTGTGAAGTGCAGAGGAATCT	5614
QY	2657	ccacctgatacagatcccaagcttlaagccgagcgcacactcccgatctggggtagaaag	2716
Db	5615	CCAACTGGATCCATGATCCCAAGTTAAGCGGCGCACCTCCGATTCGGGGTGGAAAG	5674
QY	2717	ccgaaagaactgaaatcccttaagctctgcctccatcgtcggtgttctctactctgca	2776
Db	5675	CCGAAAGAACGAAAAATCCCTTAAGGCTTGCCCTCATGCGGCTGTTCTTACTCTGTCA	5734
QY	2777	ataactctcaagtlaatgltlaaagcgcttgcagagccgaaacctccataaacctta	2836
Db	5735	ATAACTCTCAGATTATGTGATGGGATGAGACAGCGCTGAACCTCATTAACCTTA	5794
QY	2837	tctctacacttggtactactactgactccggtacaggtatataatlaacagcaactcaagg	2896
Db	5795	TCTCTCACTCGTTAATTTACTGACTCCGGCACAGGTATTAATTCACACACTCCAGAGG	5854
QY	2897	gagctcccttggggacctgtgagcctgaatlatatgctgcctctgactcaatccct	2956
Db	5855	GAGGCTCTTTAGSAACTGGGTGGCTGATCTTATAGCTTTGGCTTCAGATCACTTTATCTCT	5914
QY	2957	ggtctcaatgacagcgccacaaccccccgatgctacgtcctgcgtcttcaaggttctacgttgc	3016
Db	5915	AGTCT-----GACCTCACCCCGCAGATATCTCCATGTCACAGGATTTATGTTTGC	5965
QY	3017	ccaagacccccaataatgaagaatatgtggaatccctcaagatctccttgcagaaca	3076
Db	5966	CCAGACACCAACAAATAATGMAAAACATTTGGGAAATCCAGAGATTTCTTTTGTAAACAA	6025
QY	3077	tggaactgcataactcttaatgatgaggaatctgaaatggcagctctccagcagaagaca	3133
Db	6026	TGGAACTGTTAACCTCTAATGATGATATTGGAATGGCCAACTCTTCAGCAGATAGG	6086
QY	3137	gtaagttaactcttltgtaacaactcctcaccagttataatcaatlaatgaatgacatggg	3196
Db	6086	GTAAGTTTCTTATGTGCAACACACTATATACAGCTCTGCAACAATTTATTAACCTGCATCGG	6145
QY	3197	agatggaagaagtctggcaacagcggtgatacaaaagatgctgaagaataagaagaatgct	3254
Db	6146	ATTAAACTGTGAACCCCAAGTGTCTCTTCAGACCTTGATGATTAACCTAAATAATGAATT	6203





QY 499 ctgtgatcccggtgttggaacttaagcagcccaactctacccggttaacaaagaanaa 558  
| | | | |  
Db 4016 CTGTGATCCCGGGTTCGACACCTTAGCACCCACCTCAACCCGCTAACCAAGAAAA 4075  
| | | | |  
QY 559 ggggattctcctggcctccagacacagagcatttgaatgtatcaaaaagccgtg 618  
| | | | |  
Db 4076 GGGGAATCTCTGGGCTCCCTGACACACAGAGCATTTGATGCTATCAAAAAGGCCCTG 4135  
| | | | |  
QY 619 ctgagcagactgtcttggccctccctgaagtaactaaaccttacccttattgtgat 678  
| | | | |  
Db 4136 CTGACCGCACCTGCTCTGGCCCTCCCTGACGTAACCTTTACCTTTATGTGAT 4195  
| | | | |  
QY 679 gagcgttaaggagtagcccgagaggttttaacaaaccttaagaaacattgagagacct 738  
| | | | |  
Db 4196 GAGCTTAAGGAGTAGCCCGAGGAGTTTAACCAAAACCTTAGACCATGAGAGAGACCT 4255  
| | | | |  
QY 739 gttgctaactctgaagaagaactgtatccctgaagtagcaggtgttgccctatgtctaa 798  
| | | | |  
Db 4256 GTTGCTACTCTGTAAGAAAGCTGTATCTTACCCAGTGGTGGCCCTATGCTTAAG 4315  
| | | | |  
QY 799 gctatcgagctgttggccatacttggtaagagcgttgacaattgtactttggagcagaat 858  
| | | | |  
Db 4316 GCTATCGCAGCTGTGGCCATCTGTGCAAGAGCCTGACAAATTGACTTTGGGACAGAT 4375  
| | | | |  
QY 859 ataacttaatagcccccatactgtcatggagaactgttgcggagcccccagacagatg 918  
| | | | |  
Db 4376 ATAACTGTAATAGCCCCCAGTCATGAGAAATCGTTGGCAGCCCAAGACCGATG 4435  
| | | | |  
QY 919 atgacaaacgcccagcatgaacactatcaaaagcctgtcttcaagagaggtacattc 978  
| | | | |  
Db 4436 ATGACCAAGCCCGCATAGACCCACTATCAAAAGCTGCTTCTCACAGAGAGGTCACCTTC 4495  
| | | | |  
QY 979 gctccacacgctctcaacctgtcaacttctgtcctgaagagactga tgaacagtg 1038  
| | | | |  
Db 4496 GCTCCACACGCGCTCTCAACCCCTGCCACTTCTGCTGAAGACTGTATGAACCAAGT 4555  
| | | | |  
QY 1039 actatgatggccatcaactatgatattgagagagcttgggttcgcaagagccttaagac 1098  
| | | | |  
Db 4556 ACTATGATTTGCCATCAACTATGATGAGAGAGCTGGGTCCGCAAGGACCTTTACAGAC 4615  
| | | | |  
QY 1099 ataccgtgaactggaagaagtgtcaactgttcaactgaaggaagcagcttgtgtgtgaa 1158  
| | | | |  
Db 4616 ATACCGCTGACTGAGAAAGTGTAACTGCTTCACTGACCGAAGCACTATGCTGTGGA 4675  
| | | | |  
QY 1159 ggttaagagagatgtgtggtgcgagtggttgagcgggagcccgacagatcttggccagcag 1218  
| | | | |  
Db 4676 GGTAAAGAGATGCTGGGGGCGGTGTGTGACGGGACCCACAGATCTGGCCACAGC 4735  
| | | | |  
QY 1219 ctgcccgaagaacttaagcgcaaaaggtgagctcatatgaccttaacgcaagctttggg 1278  
| | | | |  
Db 4736 CTGCCGGAAGAACTTCAAGGCAAAAGGCTGAGCTCATGGCCCTCAAGCAAGCTTTGGCG 4795  
| | | | |  
QY 1279 ctggcgaagaagaatcatalaacttatacgaagcagagatgtacttttgactga 1338  
| | | | |  
Db 4796 CTGGCCGAAGGAATCCATAAACATTTTATACGAGACAGCATGCTTTTGGCATGCA 4855  
| | | | |  
QY 1339 caactacaggggcatctataaacaaggggtgttactactcaagcagagaggaata 1398  
| | | | |  
Db 4856 CACGTACACGGGGCCATCTATAAACAAAGGGGCTTACTTACCTTACAGAGAGGGAATA 4915  
| | | | |  
QY 1399 aagaacaagaagaatcttaagccatttagaagccttcatcttgcacaagaagctagct 1458  
| | | | |  
Db 4916 AAGAAACAAAGAAATTTCAAGCTTATTAAGAGCCTTACATTTTGGCAAAAAGGCTAGCT 4975  
| | | | |  
QY 1459 attatacaactgtcttgacatacagaagcaaaagatctatacttagaaggaagcaagatg 1518  
| | | | |  
Db 4976 ATTATACACTGCTCTGCAATCAGAAACCAAGATCTATATCTAGAGAGCAAGATG 5035  
| | | | |  
QY 1519 gctgacgggttgcagagcagcagcagcagctgttaacctctctcctaataagaagc 1578  
| | | | |  
Db 5036 GCTGACCGGGTTCCAAGCAGGAGCCAGGCTTAACCTTCTGCTTATATAGAAAGC 5095  
| | | | |

---

QY 1579 cccaaagcccaagaccagacagctacacccctagaagacttgcagagataaaagaag 1638  
| | | | |  
Db 5096 CCCAAAGCCCCAGAACCCAGACAGCAGTACACCTTGAAGAGCTGGCAAGATTAAGAAAG 5155  
| | | | |  
QY 1639 atagacagttctctgaagctcccgagggagacttgtataactcaatgtggaagaatc 1698  
| | | | |  
Db 5156 ATAGACCAAGTCTCTAGACTCCGAGGGGACCTGCTATACCTCAATATGGAGGAAATC 5215  
| | | | |  
QY 1699 ctgcccacaagaaggttlaagatalgtlcaacaacatalcagtccttaacccactaaga 1758  
| | | | |  
Db 5216 CTGCCCAACAAAGAGGTTAGAATATGTCCAAACATATCATCTGTCACCCACTAGGA 5275  
| | | | |  
QY 1759 actaaacacctgcagcagattgttgcagaacatcccttalcattgtctcgaagctacaga 1818  
| | | | |  
Db 5276 ACTAAACACCTGACGAGCTGTGTCAGAACATCCCTTATCATGTTCTGAGCTACAGGA 5335  
| | | | |  
QY 1819 gtgagctgactggtgtgcaaacattgttggccctgcagctgtgttaatgtactccctc 1878  
| | | | |  
Db 5336 GTGGCTGACTCGGTGCTAAACATTTGTGTGCTTGCACCTGTTAATGCTAATCTTTC 5395  
| | | | |  
QY 1879 agaataccctcagaagaagactaaaggagagcaaccacagcgtcacttggagttgag 1938  
| | | | |  
Db 5396 AGAATACCTCCAGGAAGAGACTTAAGGGAAGCACCAGGCGCTCAGTGGAAATGTGAC 5455  
| | | | |  
QY 1939 ttaactgaagtaaaagccggcctaataacggaacaaataltatltgttttltgtagacac 1998  
| | | | |  
Db 5456 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAATATCTATGTGTTTGTAGACACC 5515  
| | | | |  
QY 1999 tttaacgagttggtagagcttactacttaagaagaagacttcaacgttggctgaag 2058  
| | | | |  
Db 5516 TTTTCAGATGGGTAGAGGCTTATCTTCAATGAAGAAAGACTTCAACCGTGGTCTTAA 5575  
| | | | |  
QY 2059 aaaatcttgaagaaatlttccaagatttgaatlaactaaaggtlaataggltcagaacat 2118  
| | | | |  
Db 5576 AAAATCTGAGGAATTTTTCGAGATTGTGAATTAAGTATAGGTGGTACAGCAAT 5635  
| | | | |  
QY 2119 ggtcaagcttctgttgcacaggtlaagtcagagagcttgcacaagataltgggattgtatg 2178  
| | | | |  
Db 5636 GGTCAAGCTTTTGTGCTCCAGGTAGTACAGGAGCTGCGCAAGATATTTGGGATGATTGG 5695  
| | | | |  
QY 2179 aaactgaattgtgcatagaccccaagaagctcagaagagtagagaagataaataagaac 2238  
| | | | |  
Db 5696 AAACGTGATTTGTGCTACAGACCCCAAGCTTACAGACAGTGAAGATGAATGAAGACC 5755  
| | | | |  
QY 2239 attaaagaagcccttaccaaatltgaccaaagagacttgcattatgtatltgtagctc 2298  
| | | | |  
Db 5756 ATTAAAGAGACCTTACCAAAATTTGACCAAGAGATGTGCATTAATGATTGGATAGCTTC 5815  
| | | | |  
QY 2299 ctgccccttgtgtttagaggttagaagacacccctgagacaglttggcttgaacccctat 2358  
| | | | |  
Db 5816 CTGCCCTTGTGCTTTTAAAGGTTAGGAACACCCCTGAGACAGTTTGGCTGACCCCTAT 5875  
| | | | |  
QY 2359 gaattgcttaaggggagaccccccgttggcagaatltgcttltgcaataagtgat 2418  
| | | | |  
Db 5876 GAATTGCTTACGGGGGACCCCCCTGTTGATGAATTTGCTTGTATCATATGCTGAT 5935  
| | | | |  
QY 2419 gtcctgttcccaaccccttgtctctaggtcgaagggcctgaggttgggtgagcagca 2478  
| | | | |  
Db 5936 GTGCTGCTTCCACACCTGTGTTCTAGGCTCAAGAGCGCTCGAGTGGGTAGGACACGA 5995  
| | | | |  
QY 2479 gcttgaagcagctcccggaagcctactcaag ---aggagacttgcagttccacatgc 2535  
| | | | |  
Db 5996 GCGTGAAGCAGCTCCGGAGGACCTACTCAGAGAGAGAGACTTGCAAGTTCCACATCGC 6055  
| | | | |  
QY 2536 ttcaagtttgaagattcaagctatgttagagcagcagcttgcaggaanaacctcgagactcg 2595  
| | | | |  
Db 6056 TTCCAAGTGGAGATTCATGATGTTAGAGCCACCGCTGAGGAAACCTCAGACACTCGG 6115  
| | | | |  
QY 2596 ttgaaggaacttctcgtacttltgaacacacaaagcgttggaaagtgcgaagagatc 2655  
| | | | |  
Db 6116 TGGAAAGGCCCTTATCTCTGATCTTTTGAACACACCAAGGCTGTGAAGTGAAGGATC 6175  
| | | | |  
QY 2656 cccacacttgatcatalgtacccaagctlaagccggcgcacactccgaatctcggtgtgaa 2715  
| | | | |

Db 6176 TCACCTGGATTCATACATCCACGTTAGCTGGGCCACCTCCGACCTGGGGTGGAGA 6235  
 Qy 2716 gccgaagaactgaaatcccttaagcttcgctccatccgctggttccttaactctc 2775  
 Db 6236 GCCGAAAGAACTGAGAAATCCCTTTAGCTTCGCTCCATCGCCTGGTCTTACTTAC 6295  
 Qy 2776 aataactcctcaagttaagccttgtagcaagccgaactcccaataaaccctt 2835  
 Db 6296 AATAACTCCCGACGACGATGTAACGCTTATAGACACCTGAAACCCCATAGACCTTT 6355  
 Qy 2836 atctcaccctgttactactactcgcgtacaggtatataatatacaagcactcaag 2895  
 Db 6356 ATCCCTTACTGCTGCTATTATTGACCTGATACGGGTGTCACGTGTAATAGACCTGAG 6415  
 Qy 2896 ggaagctcccttggggaactgtgacgaatataatctgctccttcgaactaacc 2955  
 Db 6416 TGTTCCTCTAGAGGACCTGGGCTGACCTGATTTCTGCTCCGATTTGATTAACCC 6475  
 Qy 2956 tggctcaatgaccagcagcaccaccccgatgactcgttactcaggttactgctt 3015  
 Db 6476 CGCTGTTA-----AAAGCACACCTCCACACTAGTCCGATGTTAGGCTTATTGCTG 6529  
 Qy 3016 ccaaggaaccccaataatgaatattgtggaatccctcaggaattccttcgaaga 3075  
 Db 6530 CCCAGG---CACAGAGAAAGAGAAATACGTGGGGAATCTCTGTTGATGAGAG 6586  
 Qy 3076 atggaagctgataacttctaataatggaatggaatgcaagcttcctcgaagaag 3135  
 Db 6587 ATGAGAGCTGCTGCTACCTCCACAGCTGAGACTGGAATGGCCGATCTCTCCAGAGACG 6646  
 Qy 3136 agtaagtaactcttctgttaa 3156  
 Db 6647 GGTAAATTTCTCTTTGTCAA 6667

## RESULT 13

PEN133818

LOCUS 8750 bp DNA VRL 14-APR-2000  
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env  
 genes and LTR (class B, clone 43).

ACCESSION

AJ133818

VERSION

AJ133818.1

KEYWORDS

env gene; env protein; gag gene; gag protein; pol gene; pol protein.

SOURCE

ORGANISM

porcine endogenous retrovirus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

gene 1064..2635  
 /gene="gag"  
 mat\_peptide 1064..2635  
 /gene="gag"  
 /function="retroviral capsid"  
 /product="Gag protein"  
 join(1064..2635,2639..6223,6099..8072)  
 /note="Viral polypeptide"  
 /codon\_start=1  
 /protein\_id="CAB65341.1"  
 /db\_xref="GI:6688950"  
 /translation="MGQVTTPPLSLIDHTVRSRAHNLISVYKGGKPPNOTPCASEMP  
 TFDGWPSEGFENSEILIAVKAILIPQSPSHDQDEYTLTWDLDEDPWPKWYN  
 KPRKGPRIILAEKKNHSAEKVEPSRIYPEIEPTWPEPDPVPAQAVG  
 PSAPGAPVEGPACTRRSRGATPERDEIALLPTTYPGPGGLOLOLYWPESS  
 ADLYNKNTHNPSEDEPQRLGLVESIMESHOPMDCCOOLQTLFTEERRIILEA  
 RKNVPMKTRPTQQLQNDLMDGFLTRGMYNNAWDEGSESLIKIRALVAGLRASRP  
 TNLKVEVMQGEVPEPSVFLERLMEARFRFTPPDPTSEAKSVALATIGOSALDR  
 KIQLRLGLQAEELDLVRAEKVYIRRETEEEKERERERERERERERERERERER  
 KILAAVEGSSRREREDFRKIRSGPVSQNLGNRPDLDDQACAYCKGMAHRCNK  
 KNGKPKVLALEDKDGRRSDPLREBRVALKEGQVEFLVDTGAESHVLLQPLGL  
 KEKSWMGATGORYPWTTRRYDLVAVHTSFLYIPCPVPLGLRDLITKGAQI  
 SEBGRREVSVNKKPLTVLLDDEVLRSPOVKPPDIDLOSMLEDPQMAFEGAG  
 LAKVPPQVYLKASATPVSYROYPLSRKARBEINHPVQALLIOGILLVPOSPMPTL  
 LPVRKPTNDYRPVQDLREVKRRVODIHVPVNPYNLSALPEBNMYTVLKDADF  
 CLRLHPTISPLFAFEMRDPQGTGRLTLWRLPDGFENSPITFDEALHRLANFRIOH  
 POUTLLQYVDILLAGATKQDCEGTAKALLLESDLGASGRMLIGFATLAPLPLTE  
 LRGGORMLTEARKKTVQIDPAPTTAKQVREELGAGFCRLMIGFATLAPLPLTE  
 KGEISNAPENHOKPTDAIKKALLSAPALAPDVKRPFLVYDERGVARGVLTGIPM  
 RRPVAYLSKRLDVPASGMPVCLKIAVAAILVADAKLITLQONTIYAIPALENIYVO  
 PPRKMTNINARTHYOSLLITERVFAFPALINPTLLPETDEVTYHDCQHLIEBTG  
 VRKDLITPLTEVEVLWTEBSSYVVGKRMAGAAYVDGHTIWASSLPGTSAQAE  
 LMAITQALRLAEKGSINIYDYSRAFAVAHGAIRKQRLITSAKEIKNEEISL  
 LEALHPLRLAIHCPGHOKAKDLSGNOMADVAKOAAVALLPIETPKAPERP  
 ROYLTPEMOEIKKIDPSEPEEGTCYSYKGEILPHKEGLEVOQIHLRHLGKHLI  
 QVPTSPHVLRLPGVADSVYKRCVPQOLVANKNSRVPKRLGSGHAWMEVDFE  
 VPKAPYGNKYLVLFVDFSGWEAVPTKETSTVANKILBEIPRGRIGKIVGSDNG  
 PAPAQVSOGLAKLIGIDWKILHCAVYRROSSQVVERMKTKETLTKLTTEGLINDIAV  
 LLPELVFRVNTPGQFGLTEPELLEYGDPPLVELASVHSDVLLSOLFESRLANLEW  
 RORAMKOLREAYSGEGLQVHFREYDGVSVYRRHRAAGNLETWKGPYLLLTPTAV  
 KVEGISTMTHASHVLAIPDPGSGRAKTEMLPLRLHRLVYVNNNSPOMHTPLSM  
 RHLPTRGEPERKALPLISFASIANFLTLTROASKRLLDSSNPHPLSLTWLIDP  
 DTGVTVNSTRGVAPRGTWPELHFCILINPAKSVSTNPVRSIGFCCGTGTEKRC  
 GSESEPCRRMSCVTSNDGDKWPKISLDQRFKFSVSPKTKVMKLYDKSCSPSD  
 LDYLKISFTEKQENIQKINGMSGSIIVFYKGGAGSTGLITLRLRSTGTEPVAVG  
 GORLESLIOGAFQAINSTDPATSCMLCSSGPYEGMAKEGFVWTRKRNOCITL  
 GSRNKLTTEVSGKCTGCKGAPRSHOULCYSTVYBOASENOYLVPGVNRMNCTL  
 TPCVSSVFNSKDPFCVWQIVPVYTHPEPVYIDEDYIRNRKRPVSTIATVAVL  
 LGTAVGVTGTAALITGPOOLEKGLGELHAMTDLALAEBSVNLSESLTSSEVL  
 ONRGLDILFLREGLCALKEECGFVYDHGALRDSMSKILRELERRRRREREDOGW  
 FKGMFNRSPTMTTILSALTGLVLLILLTVGPELLINRFAVERVRSVAQIIVLRQO  
 YGSLISGEITDL"  
 2639..6223  
 /gene="pol"  
 /function="retroviral protease and polymerase"  
 /product="Pol protein"  
 2639..6223  
 /gene="pol"  
 /gene="pol"  
 6099..8069  
 /gene="env"  
 /function="retroviral envelope"  
 /product="Env protein"  
 6099..8069  
 /gene="env"  
 /gene="env"  
 8122..8750  
 /db\_xref="taxon:61673"  
 /clone="43"  
 1..629  
 /note="cap site"  
 467  
 /note="gly4"  
 primer\_bind 632..649  
 /note="gly4"

LTR  
 POLYA-signal  
 BASE COUNT  
 ORIGIN

Query Match 86.3%; Score 2864.6; DB 14; Length 8750;  
 Best Local Similarity 95.2%; Pred. No. 0;

Matches 2990; Conservative 0; Mismatches 139; Indels 12; Gaps 3;

QY 19 acaagatgctctctgctgagatatacccccactaagcaacacttttgccttgaa 78  
Db 3446 AAGATGCTCTTCTGCTGAGATTACCCCTAGCAACCCCTTTTGGCTTCGAA 3505  
QY 79 tggagagatccagtgacggaagacccgagcctcactgacccgactgcccgaag 138  
Db 3506 TGGAGATCCAGTAGGGAAGAACCCGGGAGCTCAGCTGACCCGACTGCCCAAGG 3565  
QY 139 ttcaagaactcccgacacacttttgaagagccctacacagagacttgcacattcag 198  
Db 3566 TTCAAGAAGTCCCGACCATCTTTGACGAGCCCTACACAGGAGACTGGCAACTTCAG 3625  
QY 199 atccaaacccctcagtgacccctcctcagtaagtgagtgacgcttcctggagagcc 258  
Db 3626 ATCCAAACCCCTCAGGTACCTCTCCAGTAGAGTAGAGCTGCTTCTGGCGGAGACC 3685  
QY 259 accaaacagagactgttagaagtgacgaagcactactctggaattgtcgaactagac 318  
Db 3686 ACCAAACAGAGACTGCTTAGAAGGTACGAAGGACTACTGCTGGAATTCTGACTTAGCC 3745  
QY 319 taagagagctctgtcaagaagccccagattgcagagagagagtaactatttgggtac 378  
Db 3746 TACAGAGCTCTGTAAAGAGGCCAGATTGTCAGAGAGAGGTAACTACTTGGGGTAC 3805  
QY 379 agttgcgggagcgagcagcagatgctgaagagcagcgaagaaactgtatccagata 438  
Db 3806 AGTTGCGGGGCGGAGCCAGTGGCTGACGAGGACGGAAGAAACTGTAGTCAAGATA 3865  
QY 439 ccggcccccaacacagccaaacagtgagagatttttgggagagcttgattttaga 498  
Db 3866 CCGGCCCCAACACAGCCAAACAGTAGAGATTGTTGGGAGAGCTGGATTGTTGACGA 3925  
QY 499 ctgtgagctccggggttttgagacttgcagccccactctaccgctaaccaagaanaa 558  
Db 3926 CTGTGAGTCCCGGGTTTGGACCTTAGCAGCCCACTTAACCGCTAACCAAGAAANA 3985  
QY 559 ggggagcttctcctgagctcctgaagcaccagaagactttatgctatcaaaagccctg 618  
Db 3986 GGGGAAATTCCTGAGGCTCTGAGCACCCAGAACATTTATGCTATCAAAAAGGCCCTG 4045  
QY 619 ctgagcagactgctctgagcctcctcctgacgtactaaaccccttaacccttattgt 678  
Db 4046 CTGAGCCACACTGCTGTGGCCCTCCGAGCTAACTAAACCTTTACCTTTATGTGAT 4105  
QY 679 gagcgttaaggaagtagccgaggaagttttaaaccnaaaccttaggacattggaagact 738  
Db 4106 GAGGTAAGGGAGTAGCCGAGAGATTAAACCAAACTAGACCATGAGAGAGACT 4165  
QY 739 gttgactactgtcaagaagcttgatcctgttagccaagtgttgccgtatgtgaag 798  
Db 4166 GTTGCCCTACTGTCAAAAGAAAGCTTGAATCCTGTAGCCAGTGTGGCCCTATGCTGAAG 4225  
QY 799 gctatgcagctgttgacatactgtlcaagagcgtgaacaaattgactttggagacgaat 858  
Db 4226 GCTATGCGAGCTGTGSCCATCTGTCAGAGAGCGTGAACAAATGTGACTTTGGGACGAAT 4285  
QY 859 ataacgtatagccccccatgcatgtggaacatcgttcggcagccccagacagatlg 918  
Db 4286 ATTAAGTATATAGCCCCCATGATGTGGAACATCGTTGCGACGCCCAAGCGATAGG 4345  
QY 919 atgaacaaagcccgatgaacacatacaaaagcgtcttcaacagagagaggtcaacttc 978  
Db 4346 ATGACCAAGCCCGCATGACCACTATCAAAAGCTGCTTCTACAGAGAGGGTCACTTTC 4405  
QY 979 gctccaacagcgtcttcaacacctgcaactcttctgctgaagagagactgaataacagtg 1038  
Db 4406 GCTCCACAGCGCGCTCTCAACCTCGCACTTCTTGCCGTAAGAGACTGATTAACCAAGTG 4465  
QY 1039 actaatgtatgcatacaactattgttagagagactgggttcgcaagagacttaacagac 1098  
Db 4466 ACTCAATGATTCATCACTATTTGATGAGAGAGACTGGGGTCCGCAAGAGACTTACAGAC 4525

QY 1099 ataccgctgacttgagagaagtgtctaacctggttcaactgaaggaagcaactatgttggtaa 1158  
Db 4526 ATACCGCTGACTGAGAGAGTGTACTAAGCTGTTACTACTAGCGAAGCACTATGTGTGGA 4585  
QY 1159 ggttaagagatgctgttgaggcgagtggtgaacgggaacccgcaagatcttggcgaagc 1218  
Db 4586 GGTAAAGAGATGTGTGGGGGGGGGTGTGTGAGGGGACCCACAGATCTTGCGCACAGC 4645  
QY 1219 ctgcgggaaggaacttaagcgcaaaagcgtgaactatagccctcagcaagcttggcg 1278  
Db 4646 CTGCGGGAAGGAACCTTAGCGCAAAAGCTGAGCTCATGGCCCTCAACGCAAGCTTGGCGG 4705  
QY 1279 ctggcgaaggaagatccataaacaattatagcgaacagcaggtatgcttgcactgca 1338  
Db 4706 CTGGCCGAAGGAATTCATTAACATTTATAGGAGACAGCATGCCCTTTGGACTGCA 4765  
QY 1339 caagttaacggggcactctataaacaaggggttgcttaacctcagcaggaaggaata 1398  
Db 4766 CAGGTACACGGGGCCTATTAACAAAGGGGCTGCTTAACCTACAGCAGGAGGAATA 4825  
QY 1399 aagaacaagaaggaattctcaagcctatagaagccttaactttgcaaaaagctagct 1458  
Db 4826 AAGAACAAAGAGAAATTTCTAAGCTATTAAGAGCTTACATTTGCAAAAAGCTTAGCT 4885  
QY 1459 attataactgttcttgagacatcagaagccaaagatctatatactagaggaacagatg 1518  
Db 4886 ATTATACACTGTCTGTGACATCGAATGCAAGGCAAAAGATCTCATATCTAGAGAGAACAGATG 4945  
QY 1519 gctgcagcgggttgccaagcagcagccagcgtgttaacctctgcctataatagaacg 1578  
Db 4946 GCTGACCGGGGTGTGCCAAGCAGGAGCCCGCTGTTAACTTCTGCTATATAGAAAGC 5005  
QY 1579 cccaagccccagaacccagcagcagtaacccctagaagagacttgcaagagataaanaag 1638  
Db 5006 CCCAAAGCCCCAGAAACCCAGACGACATACCTTAGAAGACTGCAAGAGATTAAGAAAG 5065  
QY 1639 atagacagcttctctgagactcctggaagggagcctgtactaactatctaggaagaatc 1698  
Db 5066 ATAGACCAATGCTGTGAGACTCCGAGGGGAGACTCTTATACCTCATATGGGAGAAATC 5125  
QY 1699 ctgcccccaagaagaggttagaatactgccaagaatatacgtctaaaccactaga 1758  
Db 5126 CTGCCCAAAAGAAAGGGGTAAATATATCTCAACAGATATCGTCAACCCACTAGGA 5185  
QY 1759 actaaacacctgcagcagtttgtaagaacatcccttaactatgtttctgaagctacaga 1818  
Db 5186 ACTTAACACTGTGACAGATTGCTGAGAACATCCCTTATATGTTGTGAGGCTACAGGA 5245  
QY 1819 gtygctgactcgtgtgtcaacatgtgtgctcctgcagcgtgtgtatgttaactcttc 1878  
Db 5246 GTGGCTGACTCGGTGTCAAAACATTTGTGTGCTGCGCAAGCTGTAAATGCTAACTTCC 5305  
QY 1879 agaataactccaagaagaagactaaggggaagccacccagcgctcaacttgggaagtggac 1938  
Db 5306 AGAGTACCTCCAGGAAGAAAGACTTAAGGGGAGCCACAGCGCTCACTGGGAAGTGGAC 5365  
QY 1939 ttcactgagtgtaaaagccggtcaaaatagcgaagaacatatactatgtgtttttagaagac 1998  
Db 5366 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATATCTATGTGTTTGTAGACACC 5425  
QY 1999 ttccaagatggttagagagcttaactcaactaagaagaagactcaacgctgtgtgcaag 2058  
Db 5426 TTTTCAGAGATGGGTAGAGGGCTTATCTACTAAGAAAGAGACTTCAACCGGTGGCTTAA 5485  
QY 2059 aaaaactcgtgaggaatattttccaagatttggaaactaagtaabaggtcagaacat 2118  
Db 5486 AAAATACTGGAGAAATTTTCCAGATTGGAATACCTAAGGTAAATCGGGTACAGCAAT 5545  
QY 2119 gttcagacttctgttcgcaagtaagtcaagagcttgcgaagatatttggagattgtatgg 2178  
Db 5546 GGTCCACTTTTGTGCCCCAGGTATGCAAGGAGCTGGCAAGATATTTGGGATGATTGG 5605

QY 2179 aaactgcatgtgcatatacagaccccccaagctcagagacaggtatagagatgaaatagacc 2238  
 Db 5606 AACTGCATGTGTCATACAGACCCCAAGCTCAGAGACAGGTAGAGAGATGAAATAAAC 5665  
 QY 2239 attaaagagaccccttaccatattgacacagagacgtgacattatgattgattgtctc 2298  
 Db 5666 ATTAAAGAGACCCCTTACCATAATGACACAGAGACTGGCATTTATGATGATAGCTCTC 5725  
 QY 2299 ctgaccccttctgcttctttaggtgtgagaaacaccccttgagacagtttgagccctat 2358  
 Db 5726 CTGCGCTTTGTGCTTTTAAAGGTTAGAAACACCCCTGAGAGATTGGGCTGACCCCTAT 5785  
 QY 2359 gaattgtctacagggagaccccccgcttgacagaaatlgctcttcacatagtgtcat 2418  
 Db 5786 GAATTGCTCTACAGGGAGACCCCGCTGTGTGAATATGCTCTGTACTACTACTGCTGAT 5845  
 QY 2419 gtgtgtgttccagagccttgtctctcctcctcagagcgtcgtgagtggtgagacaga 2478  
 Db 5846 GTGCTGCTTCCACAGCCTCTGTTCTCTAGGCTCAGAGCGCTCGAGTGGGAGCAACGA 5905  
 QY 2479 gctgtgagacagctcagggagcctactcag--aggaacttgcaatctcagctgc 2535  
 Db 5906 GCGTGAAGACAGCTCCGGGAGGCTTACTCAGAGAGAGAGACTTGCAGATTCCACATCCG 5965  
 QY 2536 ttccaagltggagattcagctcctatgttagacgcacagctgacagaaactcagactcgg 2595  
 Db 5966 TTCCAAGTGGAGATTGAGTCTATGTTAGACGGCACCGTGCAGAGAAACCTCGAGACTCG 6025  
 QY 2596 tggagagagactatctcgtacttttgacacacacagcgtgtgaaagtctgaagaaatc 2655  
 Db 6026 TGGAGAGGCGCTTATCTGACTTTTGACACACAAAGCGCTGTGAAGTCGAAAGCAATC 6085  
 QY 2656 cccacccgtgacatgacatcccaagcttgaagcgcgcacccctcagcttcagggatgaa 2715  
 Db 6086 TCCACCTGGATTCATGATCCACGTTAACTGGCCGACCTCCACATCGGGTGGAGA 6145  
 QY 2716 gccgaaaaagctgaaaaatcccttaagctcgcctccatgcgtgttcttactctgtc 2775  
 Db 6146 GCCGAAAAAGACTGAGATCCCTTAAAGCTCGCCCTCATGCGCTGCTTACTCTAAC 6205  
 QY 2776 aataactcctcaagttatgtgtaaagcctgtgtgacagcccgacactcccataaacct 2835  
 Db 6206 AATAACTCCCAAGGCGCATGTAACGCTTATAGACAGCTCAACCCCATATGACCTTT 6265  
 QY 2836 attctcaactgttactactcctacccggtacagacatataatataacagacactaaag 2895  
 Db 6266 ATCCCTTACTGCTGCTATTATGACCTGATACGGGTGACACTGTAATAAGCACTGAGG 6325  
 QY 2896 ggaagctccttgggaacctgtgtgctgaatataatgtctgccttcagctaaatccc 2955  
 Db 6326 TGTGTGCTCTAGAGGACCTGGTGGCTGATGATCTGCTCGCTCGATTGATTAAACC 6385  
 QY 2956 tggttcaatgacacagccacaccccccgatgactccgtgtccttaagggttttagcttg 3015  
 Db 6386 CGCTGTTA-----AAAGCACACCTCCCAACTAGTCCGATGATGAGGTTCTATTGCTG 6439  
 QY 3016 cccagagaccccaataatgaagaatattgtgaaatcccaagattccttgcagaaga 3075  
 Db 6440 CCCAG--CACAGAGAGAGAGAAATACTGTGGGGTCTGAGGAATCTTCTGTAGGAG 6496  
 QY 3076 atgagagctgataactcttaatgaatggaattggaaatgagcagctctcagcaagacag 3135  
 Db 6497 ATGAGAGCTGCTGCTCCTCCAAAGCTGAGAGACTGGAATAGCGGATCTCTCCAGGACCG 6556  
 QY 3136 agtaagtaactcttctgttaa 3156  
 Db 6557 GGTAATAATCTCCTTTGCTAA 6577

RESULT 14  
 LOCUS PEN293657 8763 bp DNA  
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene

and env gene, class B, clone 213.  
 AJ293657  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

env gene, class B, clone 213.  
 AJ293657  
 1. 8763  
 /organism="porcine endogenous type C retrovirus"  
 /specific host="Sus scrofa"  
 /db\_xref="taxon:70540"  
 /clone="213"  
 /cell\_line="PK15"  
 /note="class B"  
 1. 630  
 467  
 /note="cap site"  
 633. 650  
 1077. 2651  
 /gene="gag"  
 1077. 2651  
 /gene="gag"  
 /codon\_start=1  
 /product="group specific antigen"  
 /protein\_id="CAC39625.1"  
 /db\_xref="GI:14275842"  
 /translation="MGOTVTPISLTLDHWTEVRSRAHNLTVGKKGPMOTFCASEMP  
 TFDVMPSEGFENSEILAVKALIFGPGSHPODEPYLTWDLAPEPPVWKPMNP  
 KPRGPRILALGKNSHAKVSPGCIPEPEPTMPPEPOVPEPPADGAVNG  
 PSAPGAPVYBEPGAPGRSRGATPEPTDELAIPLTTPPEPMGGOLQLOLQTVSS  
 ADLNTWKTNPPESEDPORLGLVESLMSFSDPMDQDQLOLTLFTTERERITLLEA  
 RKNVADGAPTOLENIIDMGFPLTRGMDVNPAGESESLIKYQALVAGLRASRP  
 TNLAKVEVMGPNPEPSVFLERIMEAFRRFTPDPTSEAKASVALAFIGSALDIR  
 KILQRLGLOEALDLVREAEKVYRRETEEEKREKEREEREEREERDRROENKLT  
 KILAAVEGSSRERERDFKIRSGPQSGNLGNRTPLDNDQCAVCNKKHMANRNCX  
 KGNKGPVIALLEEDKD"  
 2652. 6239  
 /gene="pol"  
 <2652. 6239  
 /gene="pol"  
 /note="complete pol ORF. gag-pol is expressed as a  
 polyprotein"  
 /citation=[1]  
 /codon\_start=1  
 /product="polymerase"  
 /protein\_id="CAC39626.1"  
 /db\_xref="GI:14275843"  
 /translation="GRGSDPLPEPRITLAVEGQVPEFLVDGAHSLYLQPLGKLKE  
 KSNWKGATGROYPTWTRKTVDLAVGRVTHSLVPECPVLIGROLTKMGALQISF  
 EQGRPEVSVNKKPLTTLVLQIDDEYRLYSQVVPDDIOSMLQFQVAMETMGCLA  
 KOYPPQVIOUKASATPVSVKQYPLSRVARGIIPHYQRLIOGQILVAVOSPMPTPLP  
 VRKPGTNDYREVODLREVNRKQVODIHTVAPNPYLSALPEERNWTVLIDKAPFCL  
 RIHPTSQPLFAFERMDPGTGRGTGOLITRILPDGPKNSPFEEDALRLDLANEPHO  
 VTLLQYVDLLLAGATKQDCECTKALLLELSLUGLRASAKKQICREKVTYIGYSLR  
 GQQRWLTBARKTIVQVLPPTTAKQVREFLGTAFQCLMTLGFATLAADLYPLTKRG  
 EFSWAPRHOKAFDAIKKALLSAPALADPVTKPFTTYVDERKGVAGVAVLQTLGPMWR  
 PVALTSKKLDPVASGMPVCIAAIAVAIIVKADKLLTGNTIVIAPIHALENIVRPP

```

DRMTNARMTHTYQSLLTERTVTFAPPALNPATLPEETDEPVTHDCHOLITEGVR
KDLTDLPLTEGVALTWTPTDSSSYVEEGKKAAGVADGRTIAASSLPETSQAKELM
ALTOALRLAEGKSINLYDSRYAFATAHVGHAIYKORGLITAGSEIKKEILILE
ALHLEPKRALIHCPOGKAKADLISRONMADVADAOADAVMLPIETPKPEERRO
YTLDEMOEIKIIDOFSEPEKAGCYTSYKEILPHKEGLEYVOGIRHRTLTKHLOL
VTPSPYHVLRLPGVADSVYKHCPOLVNPANBSRIIPGKREISHPGAMWEDFTEVK
PAKYGKNTLIVYEDTFSWVEATPFRKESSTYVAKKILBEITPRGIRKVISDNGPA
EVAQVSQGLAKLIGIDMKLHCAYRPOSSQVEERMNTIKETLTKLTTETGIDMLAIL
PVLFRVNTPGOGLJTPELLYGPPLETAHSVADVLISQPLFSLKALEVRO
RAMKOLREAYSEGGLOVPHRFQVDSVYRRHHANLETFRKGYLIVLTTPYAVK
EGISTWISHASHKLAAPPDSCWRKAEKTEMLRLRLHRLVPSYNNNSPQ"
6112..8085
/gene="env"
/gene="env"
6112..8085
/gene="env"
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/translation="MHPTLSMRHLPTRGGEPRRLIPLSFASIAWFLTLTTPQASSK
RLIDSSNPHRPLSLTWLIDPDTGVTVNSTRGVAPRGTMWPELHCLRLINPAVSTP
PLIVRSYGCPTGEKEKCYGSGESFCRHSCTVSNDDGKMFISIDRKFSEFN
SGGKTKVMKLYKDKSCSPDLDYLKISPEKSGENIOKMGSMGIVFEXYKGGGA
GSTLRLRLTEGTGEPYAVGPDVLAORGPALEPRPHLPVPLTSLRPTDTPPSX
STTGILPTTPPNSPCVPYKQRLSLIQGANQAINSTDPAITSCMLCUSDGPPY
ESMAKEKGNVTRKEHNDCTGSRNKLITLVEGSGCTIGKAPSHOHKESTVVEY
ASENDLVGVNRMWAKNDGLPCVSTSVNOSKGCQVQVPRVYHPEVVDLDEY
DYRNPKEPEVSLTAVMLGTAAGVOTGALITGPQOLEKELGFDHSMGATDLR
ALEESYNSLESLTSEVILONRGDLTLFREGGIDAAKKECCFVLVHSGAIRD
MSKIRLERRRERREDOGWPEGWFNPSPMWTTLTSLTGLVYLLLLITVGPCLIN
RFLVAVRERVSVAQIMVLRQYGLISQETDL"
8133..8763
LTR
8655..8660
POLY_signal
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
Query Match 86.3%; Score 2864.6; DB 14; Length 8763;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2990; Conservative 0; Mismatches 139; Indels 12; Gaps 3;
QY 19 acagatgacctctctctgacctgagattacacccaccagcaaacacttttctgctcgaa 78
Db 3459 AAGAGATGCTCTTCTTCTGCTGAGATTACACCCACCTAGCCAACTTTTGGCTTTCGAA 3518
QY 79 tggagagatccaggtacggaagacccgagctacacttgagaccgagctgcccagaag 138
Db 3519 TGGAGAGATCCAGGTACGGGAAGAACCCGGGACCTGACCTGGAGCCGACCTGCCCCAAGG 3578
QY 139 ttcaagaactccccaccatctttgacgaagccctacagaaggaacttgcccaactcaag 198
Db 3579 TTCAAGAACTCCCGCACTTTTGAAGAGCCCTACACAGGAGACTGGCCCAACTTCAAG 3638
QY 199 atccaacaacctcaggtgacctcctccagtaagctgagctgcttttggcggagagcc 258
Db 3639 ATCCAAACACCTCAGGTGAGACCTCTCCAGTACGTGATGACCTGTTTGGCGGGAGCC 3698
QY 259 accaacaagagctgcttagaaggtacgaagcaactgctgtaattgctcctaagc 318
Db 3699 ACCAAACAGAGACTGCTTAGAGGTACGAAGGCACTACTGCTGGAATTTGCTGACCTAGGC 3758
QY 319 tacagaacctctgcttagaagagccagatttgagaagagaaggttaacataacttgggtac 378
Db 3759 TACAGAGCTCTGCTAAGAGGCCAGATTTGAGAGAGAGAGGTAACTACTTGGGTGAC 3818
QY 379 agtttcgagggcgagcgatgctgacgagagcagaagaagaaactgattccagata 438
Db 3819 AGTTTGGGGGCGGCGAGGATGCTGACGAGAGCGACGGAAGAAACTTATGATCCAGATA 3878
QY 439 cgggccccacaacacgaagaagttagagagtttttggggacagctggaatttgcaga 498
Db 3879 CCGGCCCAACACGACCCAAACAGTGAAGAGATTTTGGGGAACAGCTGATTTGGACA 3938

```

```

QY 499 ctgtgagtcgggggttgagacacttaagagcccaacttaaccgcttaacaaagaanaa 558
Db 3939 CTGTGAGTCCCGGGGTTTGGCAGCTTACAGAGCCCACTACCCGCTTAACCAAGAAAAA 3998
QY 559 ggggagatctctgggacctgaagacccaagagatttgatcatataaaagccctg 618
Db 3999 GGGGAATTCCTCTGGGCTCTGAGCACCAGAGGATTTGATGCTATCAAAAAGCCCTG 4058
QY 619 ctgaagcgaacctgaccttgccctcctgaagtaactaaaccccttaacccttatgtgat 678
Db 4059 CTGAGCGACACCGCTGCTGCGCCCTCCCTACAGTAACCTTAACCTTTATGTGAT 4118
QY 679 gagcgtaaggagtagtcccgagagtttaacccaacaccttagaagcaatggaagagacct 738
Db 4119 GAGCGTAAAGGAGTAGCCCGGAGAGATTTTAACCCAAACCTTAGAGCATGAGAGAGACCT 4178
QY 739 gttgacctctgctaaagaagcttgaatcttgtagcaatggttggccgcttgctgaag 798
Db 4179 GTTGGCTACTGCTCAAGAAAGCTTGATCTGTAGCCAGTGTGGCCCGTATGCTGTGAAG 4238
QY 799 gctatcgagctgtgacatactggtcaagagacgtgacaaatgtaactttgagagaagat 858
Db 4239 GCTATCGCAGCTGTGCGCATACTGTTCAAGAGCGTGACAAATTTGACTTTGGAGACAGAT 4298
QY 859 ataactgtaataagcccccaatgcatltgagaacaatcgltggcgagcccccaagcagatbg 918
Db 4299 ATAACTGTAATAGCCCCCATATGATGAGAAACATCTGTCGGAGCCCGCCAGACGATAG 4358
QY 919 atgacaaagcccgatataccactataaagccgtcttccaaagaagagtgtaacttcc 978
Db 4359 ATGACAAAGCCCGCATACCCACTATCAAGACCTGCTTCTCAAGAGAGGCTCACTTTC 4418
QY 979 gctccacagcgcgtctctaacaccttgcaactctctctgctgaagagactgataaaccagtg 1038
Db 4419 GCTCCACAGCGCGCTCTCAACCCCTGCACTCTTCTGCTGAAGAGACTGATGAACAGTG 4478
QY 1039 actatgattgcatcaactatgtattgaagagagctggggtccgcaagagaccttaacagac 1098
Db 4479 ACTCATGATTGCGCAACAATATTGATTGAGAGAGCTGGGGTCCCAAGGACCTTACAGAC 4538
QY 1099 ataccgctgactggaagagttctaaccctggtcactgaacggaagcagctatgtgtgaa 1158
Db 4539 ATACCGCTGACTGGAGAGTGTCTAACCCTGTTCACTGACGAGAACAGCTATGTGTGGA 4598
QY 1159 ggtlaagagatggtctgggagcgagtggttggacggagacccgacagatctgggacagac 1218
Db 4599 GGTAAAGAGATGCGTGGAGCGGCGGTGTGAGACGGAGACCCACAGATCTGGCGCAGCAGC 4658
QY 1219 ctgcccgaagagacttgaagcgaagaagctgaagctcaatgagccctcaagcagcttggcg 1278
Db 4659 CTGCCGGAAGGAATCTCAGCCCAAAAGGCTGAGCTCATGAGCCCTACGCAAGCTTTGCG 4718
QY 1279 ctggcgaaggaataatcataaacaatlatalacgagacagatgtgctttgagacgtca 1338
Db 4719 CTGGCGGAAGGAATTCATTAACATTTATACGAGACAGAGTATGCTTTGGCAGCTGCA 4778
QY 1339 caagtlacaggggacatctataaacaagaaggggttcttaacctcaagcagggaggaata 1398
Db 4779 CACGTACACGGGGCATCTATTAACAAAGGGGTTGCTTACTCAGCAGGAGGAGGAATA 4838
QY 1399 aagaacaagaaggaattctaagctatgaagccttaacattggccaanaaagctgct 1458
Db 4839 AAGAACAAAGAGGAATTTCTAAGCCTATTAAACCTTTTACCTTTTGGCAAAAAGGCTGAGCT 4898
QY 1459 attatacaactgctctgagacacaaagacaaagatcatalatctagaagggaacagatg 1518
Db 4899 ATTATACACTTCTCGATGACATCAGAAAGCCAAATCTCATATCTAAGGAGCAACCAATG 4958
QY 1519 gctgacgggttgcgaagcagggcagccagtgcttlaaccttgcctataatagaagac 1578
Db 4959 GCTGACCGGTTGGCAAGCAAGCAGCAGCGCTTAACTTTCCTATTAATGAGAAACG 5018
QY 1579 cccaagaccacgaagaccagacagtlacacaccttaagaagactggtgcaagagataaanaag 1638

```

```

|||||
Db 5019 CCCAAGACCCAGAACCCAGAGACAGTACACCTAGAGACTGGCAAGATATAAAG 5078
QY 1639 atagacaaattctttagagctccggagggagactctataactatagtggagaatc 1698
Db 5079 ATAGACCAATCTCTGAGCTCCGGAGGGACCTGCTATACCTCATATGGAGGAATATC 5138
QY 1699 ctcccccaaaagaaggttagaataatgtccacagatatacgtctaaacccactaga 1758
Db 5139 CTGCCCCACAAAGAAAGGTTAGATATGTCCAACAGATATCATGCTCCACCCTAGGA 5198
QY 1759 actaaacacctgacagagtttggtaacaaatccctatcatgtcttgaagctacaga 1818
Db 5199 ACTAAACACCTGACGAGCTTGGTCAAGAACATCCCTTATCATGTCTGAGCTACAGGA 5258
QY 1819 gtgagctagctgttggtaacaaatgtgtcgcctgcagctgtttaaagtatactctcc 1878
Db 5259 GTGGCTGACTCGGTGGTCAAAACATGTGTGCTTGCACCTGGTTAATCTAATCTTCC 5318
QY 1879 agaatccctccagaagaagactaaagggaagccacccagcgctcaactgtgaagtgaac 1938
Db 5319 AGAATACCTCCAGAAAGAGACTAAAGGGAGACCCACCGGGCTCAGTGGAGATGGAC 5378
QY 1939 ttaactgaagtaaacgcggctaaatacggaaacaaatatctatgtgttttgaacac 1998
Db 5379 TTCACAGAGGTAAAGCCGGCTTAATACGAAACAAATATCTATGTGTTTGTAGACACC 5438
QY 1999 tttagagatggtagagcttaactactaagaagaagacttaacggtgtgtgaag 2058
Db 5439 TTTTCAGAGTGGTAGAGCTTATCTACTAAGAAAGACACTTCAACCTGGTGGCTTAA 5498
QY 2059 aaatactgagagaaattcttccaagatttgaaatacctaaggtaataaggtcagacat 2118
Db 5499 AAAATCTGGAGGAATTTTTCAGATTTTGGATATACCTAAGGTATCGCGTACAGACAT 5558
QY 2119 ggtcagcttgcgttgccaggtlaagtcaggagacttggccaagatatgtggatgtatgg 2178
Db 5559 GGTCACGCTTTTGTGCTCCAGAGTAATCAGGAGACTGGCCAAATATTTGGGGATTATGG 5618
QY 2179 aaactgatttgcatacagaagcccaaaactcagagcaggttagaaggttagaagac 2238
Db 5619 AAAGCTCATTTGTCATACAGACCCCAAGCTCAGACAGTGAAGAGATGATGAAAC 5678
QY 2239 ataaagaagaccttaaccaatlgacacacagagacttgatattatgtatgtatgtctc 2298
Db 5679 ATTAAGAGAGCCCTTACCAATTTGACCAAGACAGAGCTGATTAATGATGTGATCTGC 5738
QY 2299 ctgaccttgcgtcttttaagggtgaggaacacccctgagcagtttgggtgacccctat 2358
Db 5739 CTGCCCTTGTGTGCTTTTGGGTTAGGAACACCCCTGACAGCTTGGGCTGACCCCTAT 5798
QY 2359 gaaatcttaaggggagaccccccgcttgacagaatgcttggacaatgtgtat 2418
Db 5799 GAATTCCTCTACGGGGGACCCCGCTGTGTAAGAAATGCTTCTGTATATGTCTGAT 5858
QY 2419 gtgctgttccacagcttgccttctcctaagctcaagcgctgagtggtgaggaagca 2478
Db 5859 GTGCTGCTTCCAGCGCTTGTCTCTAGGCTCAAGGCGCTCGAGTGGTGAAGCAACA 5918
QY 2479 gegtgaagcagctccggagagccttaacg---agagacttgcgaattccacatgc 2535
Db 5919 GCCTGGAACAGCTCCGGGAGGCGTACTCAGAGAGAAGGAGACTTCAAGTTCCACATGCG 5978
QY 2536 ttccaagttgagatcaatcaatgtatgttagagccacgctgacggaagaaactgagaactgg 2595
Db 5979 TTCCAAGTGGAGATTTCAGCTATGTGTAAGCGCACCATGCGAGGAACCTCGAGACTCGG 6038
QY 2596 tgaagaagacatctatctgactttagcacacacaaagcgctgtaaaatcgaagaatc 2655
Db 6039 TGGAAAGGCGCTTATCTGACTTTTGGACACACCAAGCGCTGTGAATTCGAAGGAATC 6098
QY 2656 cccacctgatatcatgcatccacglttaagccggcgcaactcccgatcggggtgga 2715
|||||

```

```

Db 6099 TCCACTGGATCATGCATCCACCTTAAGCTGGCGCCACCTCCCGACTGGGGTGGAGA 6158
QY 2716 gccgaagaagactgaataatcccccttaagcttgcctccatcagcggtgtcttaactgtc 2775
Db 6159 GCCGAAGAGCTGAGATCCCTTAAAGCTTGCGCTCCATGCTCGTGTCTTACTCTTAAC 6218
QY 2776 aatacctccaaatgaatgtaaacgcctgttgaacagccggaactcccaataaccc 2835
Db 6219 AATAACTCCAGCGAGTACTAATAAGCGCTTATGACAGCTGCAACCCCATATGACCTTT 6278
QY 2836 atctcactcgttactactagctactccgtaacaggtatataataataacagcactaagg 2895
Db 6279 ATCCCTTACCTGGCTGATTAATGACCTGATACCGGTGCTCACTGTAATATGACCTCGAG 6338
QY 2896 ggaagctcccttgggagacactgtgtcctgaatataatgtcgtccttcgaatgaatcc 2955
Db 6339 TGTGCTCTCTAGAGGACCTGTGTGCTGAGACTGATTTTGTGCTCCGATTTAATACC 6398
QY 2956 tggctcaatgacagagcacaccccccgaatgtaactcgtgtcttaaggtttagt 3015
Db 6399 CGCTGTAT-----AAAGCACACCTCCACACCTAGTCCGTAATGAGGTTCTATGCTG 6452
QY 3016 cccaggaaccccaataatgagaatattgtggaatccatccatgatttcttgaagca 3075
Db 6453 CCCAGG---CACAGAGAAAGAAATACTGTGGGGGTTGTGGGAATCTTCTGTAGAG 6509
QY 3076 atgagctgcatcaatcctaataatggaatggaaatggcagctctcgaagaagc 3135
Db 6510 ATGAGCTGCTTACCTTCAACGATGAGACTGGAAATGGCCGATCTCTCCAGGACCG 6599
QY 3136 aglaagtaactcttgttaa 3156
Db 6570 GGTAAATTTCTCTTTGCA 6590

```

```

RESULT 15
PERY17013
LOCUS Porycine endogenous retrovirus gag, pol, env genes. 14-APR-2000
DEFINITION Porycine endogenous retrovirus gag, pol, env genes.
ACCESSION Y17013
VERSION Y17013.1 GI:3005985
KEYWORDS env gene; gag gene; pol gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus.
Virus: Retroviridae; Retrovirinae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 7808)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retrovirus replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 7808)
AUTHORS Tonjes, R.R.
TITLE Direct Submision
JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG
FEATURES
Source
1..7808
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
260..1834
/gene="gag"
260..1834
/gene="gag"
/codon_start=1
/protein_id="CAI6581.1"
/db_xref="GI:3005985"
/db_xref="SPTREML:073504"
/translation="MGCTVTPLSLIDHWTEVRSRAHNLISVQYKKGPMTFCASMP
TFDVGWSEGTFSNSELIAVKAIIFQTPSSHPDDEYILTWDLAEDPPWPKWMLN

```



KPRKPRILLALGKRNKHSKEVPSRITYPIETEEPTWPEQVPVPPYPAQAVRG  
 PSAPGAPVVEGPAAGTRSRGATPERTDIALITRTGPPMPGGOLOPLQWPFSS  
 ADLYNMKTNPSPESDPOBLTGVLVESLHPTDCCOOLQTLTTERREHLEEA  
 RKNVPGADRPLOLONIDMGFLTRGMYNYNAEGESIKITRYKOLVIGLRASAPR  
 TNLAKVEWQGNPPSPVFLERLMAFRFPDPSVOKASVALATPGIADLR  
 KTIORLEGOAEALRDLYRAEKAYIGRETEEEKEDRKEREERERDRKOKULT  
 KIILAAVEGSSREBERFRKIRSGPROSNGLRNPLDKDOCAVCYCKEKGHWRNCPK  
 KGNKGPVLALEEDKD  
 2120..4726  
 /gene="pol"  
 /codon\_start=1  
 /protein\_id="CAA76582.1"  
 /db\_xref="GI:3005987"  
 /translation="MGASISFEQGREPVSNKPIVLTLOLIDDEYRLSPVKPPD  
 IOSMLDFOPQMAETAGMGLAKOVPROVIOUKASATPYSVROYPLSRERREGIWHVQ  
 RLIOGGLVPSPPWNTPLPVKPGSTINDIRVODLREKNRVDIHTVPPNPILS  
 ALPERKNWTVLVDKDAFCPLRLHTPSQPLFAEMHDPEVGTGQLTWRLPOGKNS  
 PTJEDALHDLNFRIOHPVTLQYVDLLAGATKODCEGTKALLLESDLSGR  
 ASAKKAOICREVTYLYSLRGORMLTARAKTVOIIPAPTAOKVRELFAGSCR  
 LMIPGATLAPLYPLTKEKEFSNAPEHOKAPDAIKALLSAPALAPDVMKPLPY  
 VDERKGVAVLQTLGPMRRPVAYLSKTLDPVASMPCUKAIAVAALLVYDADKLT  
 LGONTIVAPHALENTYROPDDRMNTNMTHTOSLLTERVFEAPALNPAATLPE  
 EIDEPVTHDQHLIETGVKRDLDIPLTEGLVMTFTDSSVVEGKMAAAYVDG  
 THTIMASSLEPSTSAQKAEMLALQALRLAEGSINIYDSRYAFATAHGAITYOR  
 GLTASGREIKNEELISLEALHLRLAIIHCPGHOKADLISGNOMARVAKOA  
 AQAVNLPIETPRAPEROYLTEDMOELKKIDOFSTPEGCTSYCKELLPKES  
 LEYVOOIHRLHGTGKLOIVRTSPYHLRLPGVADSVKRCYCOLNANPSIHP  
 GRLRSGHPGAMHEVDFTEVKRAKYNKLLVPLVDFTSQWVAEYPLTKETIVAKKY  
 MKRFFRDLXLR  
 5294..7267  
 /gene="env"  
 /codon\_start=1  
 /protein\_id="CAA76583.1"  
 /db\_xref="GI:3005988"  
 /translation="MHPTLSMRHLPTRGGEPRRLRIPLSFASIAFWLTTLTPQASAK  
 RLIDSNPRLSLTWLIDPDTGVNSTRGVAPRGTWMPPLHPCRLINAVASTP  
 PNLVRSYGFCCPQTEKEKYGSGSESPFRSCVSTNGDMKMPISLORDRKEFPV  
 SGGRKXVMKLYIKDKSCSPSDIYLKISTEKGKQENTOKMNGHSMGIVFKTGGV  
 GSTLTIRLLETETEPVAVGPKVLAEGPAPLEPVLVPLQTLSPDLTTPSSN  
 GTGLIPLTPRNSPGVPKTSQRLFSLLQGAFOAINSDPATSCMCLSSGPPY  
 EGNMKGKENVTKEHNRDCTGSRNKLTLTEVSGGTGICGAPSHOHLCTSYVEEO  
 ASENOLVPGYNNMMACTGLTPCVSSSVENOSKDFCVAVOIVPYYHYHBEVVDL  
 DYRNRPKREPSLTSLAVMLGTAAGVGTGTGALITGPOOLEKELGELHAMTEDLR  
 ALSEYSNLESLTSLSEVYONRGDILLPLREBGLCALALEBCCFTYVDHSGAIRDS  
 MSLRRLERRRRRERADOGMEFGNPNRSPWMTTLLSLTGPLVLLLLTVGPCLIN  
 KEVAFVRENVGAVOIMLVLRKQYOGILSDGETDL"  
 BASE COUNT 2077 a 1971 c 2047 g 1713 t  
 ORIGIN  
 Query Match 86.1%; Score 2857.4; DB 14; Length 7808;  
 Best Local Similarity 95.3%; Pred. No. 0;  
 Matches 2992; Conservative 0; Mismatches 136; Indels 13; Gaps 4;

2822 ATCCAAACCCCTCAGGTGACCCCTCTCCAGTACGTGATGACCTGCTTGGCGGAGGC 2881  
 259 accaaacagagctgcttaagaagtaagaaggaacactctctgtaattctctactaagc 318  
 2882 ACCAAACAGAGCTGTTAAGAAGTACGAAGGACACTCTGGAATTTGCTACCTAGGC 2941  
 319 taagaagctctgtaagaagcccaagatttcaagaagaggttaacataacttggggatc 378  
 2942 TACAGAGCTCTGCTAAGAAAGCCCGAGATTTCAGAGAGAGAGTACATCTTGGGGTAC 3001  
 379 aglttgcggggcgagcaagctgctgacgagcaagcaagcaagcaagcaagcaagca 438  
 3002 AGTTTGCAGGGGGGCGGAGGATGCTGACGAGGACAGGAAAGAAACTGATCCAGATA 3061  
 439 ccggccccaacacacgcaaaacaaagtggagaggttttggggacagctggaatttgcaga 498  
 3062 CCGGCCCAACACACACCAACAAAGTAGAGAGTTTGGGGACAGCTGGATTTTGCAGA 3121  
 499 ctgtgataccggggtttgcagacctagcagcccaactctaccgcttaaccaagaagaaa 558  
 3122 CTGTGATCCGGGGGTTTCCGACCTTAGAGCCCACTTACCCCTTAACCAAGAAAAA 3181  
 559 gggggatctctctggtctctgagcaccagaaggaatctgatacaaaaaaggccctg 618  
 3182 GGGGAATTCCTCGGGCTCTGAGCAGCAGAGGCAATTGATGATCAAAAAAGGCCCTG 3241  
 619 ctgaagagacccgctctgctgcccctctcctgaagaaacttaaccttaaccttatgagat 678  
 3242 CTGAGGCACTGCTCTGCGCTCTCCCTCGTACGTAACCTTAACCTTTATCTGAT 3301  
 679 gagcgtaaggagtagcccgagaggttttaaccacaacccctagaaccaatgagagacct 738  
 3302 GAGCGTAAGGAGTAGCCCGAGAGCTTTAAACCAACCTTAGAGCACTTGAGAGAGACT 3361  
 739 gttgcctactgtcaagaagcttgatctctgtaagcagtggttggccgtatgtctgaag 798  
 3362 GTTGCTACTCTGTCAAGAAGAGCTTGATCTGAGCAAGTGGTGGCCGCTGCTGAG 3421  
 799 gctacgaagcctgtgccaactatgtgtcaagaagctggaacttgaacttggagagat 858  
 3422 GCTATCGACGCTGTGGCCATCTACTGTCAAGACGCTGCACAAATTTGACTTGGAGAGAT 3481  
 859 ataactgtaataagcccccaatgcatltggaagacatcgtctgcgagcccccaagacgatg 918  
 3482 ATAACGTGAATAGCCCCCATGATGGAAGAATCGTTGGCAGCCCCAGACGATGG 3541  
 919 atgaccacagcccgatgaaccacatacaagaagctgcttctcaagaagaggttcaacttc 978  
 3542 ATGACCAACGCCCGCATGACCCACTATCAAAAGCCGTGTTCCACAGAGAGGTCACATTTC 3601  
 979 gctcaccaagccgctctcaacccctgcaactctctgcttgaagaagcagatgaaccagt 1038  
 3602 GCTCCACCAAGCCGCTCTCAACCCCTGCACTCTGCTGTAAGAGACTGATGAACCAAGTG 3661  
 1039 actcaatgtatgcatcaactatltgattgaagaagactgggttcgcaagaagacttaccagac 1098  
 3662 ACTCATGATTTGCATCAACTATGATTGAGAGAGACTGGGGGCCCAAGAGACTTACAGAC 3721  
 1099 ataccgctgactggagaagtgtctaactcgttcaacttgaacggaacagctatgtgtgaa 1158  
 3722 ATACCGCTGACTGGAGAGTGTCTAACCCTGTTCACTGACGAGAACACTATGTTGTGAA 3781  
 1159 gtaagaagatgctgctggggcgagtggtgaagaggaacccgacagatcttggcgagcagc 1218  
 3782 GGTAAAGAGATGCTGGGGCGGCGGTGTGAGACGGGACCCACACGATCTGGGCCACGAC 3841  
 1219 ctgcccgaaggaacttcaagcaaaagcttgaagctatgagccctcaagcaagcttgcg 1278  
 3842 CTGCCGGAAGGAACCTTACGCCAAAAGGCTGAGCTCATGGCCCTACAGCAAGCTTTGCGG 3901  
 1279 ctggcggaaggaatcatataaacttatacgggaagcaggtatgcttcttgcagtga 1338

Db 3902 CTGCGCCGAGGGAATCCATTAACATTATACGACAGCAGTATGCTTTGCGACTGCA 3961  
QY 1339 cagctacacggggccatctataaacaaggggtgtgtcttaacctcagcaggaggaata 1398  
Db 3962 CACGTACACGGGGCCATCTATAAACAAGGGGGTGGCTTACCTACGACGAGGAGGAATA 4021  
QY 1339 aagaacaagaagaatcttaagcctataggaagccttaacatttgcacaaaggtact 1458  
Db 4022 AAGAACAAAGAGGAATTTCTAAGCTTATAGAACCTTACATTGTCACAAAAGGCTAGCT 4081  
QY 1459 attatacctgtcttgacatcagaagaagccaagaatctctatctataggaagccagatg 1518  
Db 4082 ATTATACACTGTCTTGACATCAGAAAGCCAAAGATCTCATATCTAAGGGAACCAAGT 4141  
QY 1519 gctgaccggtgtgccaagcagcagccagctgttaacctctgtcctcaatagaaag 1578  
Db 4142 GCTGACCGGGTGGCCAAAGCAGGCGCCAGGCTTTAACTTGTGCTTATATATAGAAACG 4201  
QY 1579 cccaagccccagaaccagaagcagctacacccctagaagacttggcaagataaanaag 1638  
Db 4202 CCCAGAGCCCGCAGAACCCAGACGACAGTACACCTAGAAAGCTGGCAAGATTAATAAAG 4261  
QY 1639 atagaaccaattcttgagactcggagagagactgtctataccctcatatgtggaagaaac 1698  
Db 4262 ATGATACCAAGTCTCTGAGACTCCGAGGGAGCTGTATACCTCATATGGAAGGAATC 4321  
QY 1699 ctgccccacaagaaggttagaataatgtccaacagatacatcgtcttaaccacccactaga 1758  
Db 4322 CTGCCCCACAAAGAGGGTTAGAAATATGTCACACAGTACATGCTCTCACCCACCTTAGGA 4381  
QY 1759 actaacaacctgcagcagctgtgtcaagaacalcccctcatatgttctgaagctaccaga 1818  
Db 4382 ACTAAACACCTGCAGAGTGTGTGAGAACATCCCTTATCATGTCTGAGAGTACCAAGA 4441  
QY 1819 gtggctacactcgtgtgtcaaacattgtgtcctgtcagctgttgaatgaactccctcc 1878  
Db 4442 GTGGCTACCTGCGTGGTCAACATGTGTGCTGCTGCGAGTGGTATGCTATTCCTTCC 4501  
QY 1879 agaatacctcagaagaagactaagaagagagacacccagcgctcaacttggagagagag 1938  
Db 4502 AGATATACCTCCAGGAAGAAGACATAAGGGAAGCACCACCGCCTACCTGGGAAGTGGAC 4561  
QY 1939 ttaactgaagtaaaagccggttaataatcggaaacaaatctatctgttcttgaagacac 1998  
Db 4562 TTCTACGAGGTAAAGCCGGCTTAATATCGGAACAAATATCTATTGTTTGTGAGACAC 4621  
QY 1999 ttctcaagatgggttagaggttactactcaagaagaagactcaacccgtgtgtgctaaag 2058  
Db 4622 TTTTCAGGATGGGTAGAGGCTTATCTACTTAAAGAAAGACCTTCAACCGTGGTGGCTAA - 4680  
QY 2059 aaaaacttgagaagaattttctcaagatttggaataccttaaggttaaggttagagacat 2118  
Db 4681 AAAATACTGAGGAATTTTCCGAGATTGGAATACCTTAGGTATTCGGGTGACACAT 4740  
QY 2119 gttccagcttctgtcccaagtaagtcagggacttggccaagatalttgggaattgattg 2178  
Db 4741 GGTCCAGCTTTTGTGCCCCAGGTAAAGTCAAGGACTGCCAAGATATTGGGGAATTGATTG 4800  
QY 2179 aaactgcattgtgcatatacagaccccaagctcaagtcaggttagagaggttagaataacc 2238  
Db 4801 AAACCTGATTGTGATACAGACCCCAACCTCGAGACAGGTAGAGATGAATAGAACCC 4860  
QY 2239 attaaagagaccccttaacaaatgacacagagagactgcatatagattgtgtctc 2298  
Db 4861 ATTAAAGAGACCTTACCAATTTGACACAGAGACTGGCATTAATGATTGATAGCTCTC 4920  
QY 2299 ctgaccttgtgcttctttaggggttagaagaacacccctggaacgttgggtgacccctat 2358  
Db 4921 CTGCCCCCTTTGTGCTTTTAAAGGTTAGGAACACCCCTGGAAGTTGGGCTGACCCCTAT 4980  
QY 2359 gaattgcttaagggagacccccccgttggcagaagattgaccttgcacatagtgtat 2418  
Db 4981 GAATTGCTCTACGGGGGACCCCCCGTTGTGTAAGAAATTGCTTGTATACATAGTGTGAT 5040

QY 2419 gtgctgttcccaagccttgttctctagagctcaagaagcgctcaggtgggttagagacga 2478  
Db 5041 GTCTCTCTTTCCAGGCTCTGTCTTACGGCTCAAGGCCCTGAGTGGGTGAGGACGA 5100  
QY 2479 gcttgaagaagctccgggaagcctactcag - - - aggaacttgcagattccacatcgc 2535  
Db 5101 GCGTGAACAGAGCTCCGGGAGGCTACTCAGGGAAGAGAGCTTGCAGATTCACATGCG 5160  
QY 2536 ttccaagttgggattcagltcatatgttgaagcagcgttgcaggaacctcgaagactcgg 2595  
Db 5161 TTCCAAGTGGGATTCAGTCAATGTATGACGCCACCGCAGGAGAAACCTCGAGACTCG 5220  
QY 2596 tgaagaagacattatctgtcttacttggacacaccaaagcggttgaagtcgaagaatc 2655  
Db 5221 TGAAGAGGCCCTTATCTGCTACTTTGACACACACACAGGCTGTGAAATCGAAGGAATC 5280  
QY 2656 cccaacttgatccatgcatcccaagltlaagccggcgccacactccagattcgggttgaag 2715  
Db 5281 TCCACCTGATTCATGTCATCCACGTTAAGCTGGCGCCACCTCCGACTCGGGGTGAGA 5340  
QY 2716 gccgaagaagctggaataatcccttaagcttgcctccatcgcgtgttcttactctgtc 2775  
Db 5341 GCGGAAGAAAGCTGAGAAATCCCTTAAAGCTTCGCTCCATCGCCTGTTCTTACTTAAC 5400  
QY 2776 aataactcctcaagtttaagtaaacgcttggagagccgaactccataaacctt 2835  
Db 5401 AATTAATCCCGCAGCCAGTAGTAAGGCTTTAAGACGCTTGAAACCCCATAGACCTTT 5460  
QY 2836 atctcacccttgttacttactcagctccggtacaggtlaataatlaaacagacccaag 2895  
Db 5461 ATCCCTTACTTGCGTATTTAGACCTGATMGSGGTGTCACTGTAATAGACCTGAGG 5520  
QY 2896 ggaagctcccttggggaccttggctggaattatatagtctgtccttgcataagtaacc 2955  
Db 5521 TGTTCCTCTAGAGGACACTGCTGGCTGAACTGATTTCTGCTCCGATTGATTAACCC 5580  
QY 2956 tggctcactgaagccagccacacccccagatgtactcgttgcgttgaaggttttagttg 3015  
Db 5581 CGCTGTTA - - - - - AAAGCACACCTCCCAACTTACCTGATTTATGGGTCTATGCTG 5634  
QY 3016 cccaagaccccccaataatgaagaatattgttgaataatcctcaggaatttcccttgaagca 3075  
Db 5635 CCCAGG - - - - - CACAGAGAAAGAGAAATFACMGGGGGTTCGTAGGAATCTTCTGAGGAG 5691  
QY 3076 atggaagctgcaataacttcaatgagtgggaattggaatggccagctctcagcaagacag 3135  
Db 5692 ATGGAGCTGCGTCACTCCAAAGATGAGAGACTGGAATGGCCGATCTCTCCAGGAGCG 5751  
QY 3136 agtaagttactccttgtttaa 3156  
Db 5752 GGTAAATTTCTCTTTGTCAA 5772

Search completed: February 24, 2002, 02:33:44  
Job time: 19935 sec



## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8209	100.0	8309	6	A66553	A66553 Sequence 3
2	8171.4	99.5	8196	6	A66552	A66552 Sequence 2
3	8039.8	97.9	8763	14	PEN293657	AJ293657 Porcine e
4	8011.6	97.6	8750	14	PEN133818	AJ133818 Porcine e
5	7965.6	97.0	8918	14	PEN133816	AJ133816 Porcine e
6	7644	93.1	7808	14	PER17013	Y17013 Porcine end
7	7197.2	87.7	8849	14	PEN133817	AJ133817 Porcine e
8	6493.2	79.1	8918	14	PEN293656	AJ293656 Porcine e
9	6282.4	76.5	8132	6	AR130475	AR130475 Sequence
10	6279.2	76.5	8132	4	AF038690	AF038690 Sus scrof
11	6207.8	75.6	8132	4	AF038599	AF038599 Sus scrof
12	5628.8	68.7	7873	6	AX052638	AX052638 Sequence
13	5528.2	67.3	7362	6	AX052634	AX052634 Sequence
14	5489.8	66.9	7333	6	AF038601	AF038601 Sus scrof
15	5489.8	66.9	7333	6	AR130474	AR130474 Sequence
16	5025.8	61.2	6076	6	AX052636	AX052636 Sequence
17	5025.8	61.2	6076	6	AX052647	AX052647 Sequence
18	4986	60.7	8060	6	AR130473	AR130473 Sequence
19	3755.4	45.7	4402	6	AX052635	AX052635 Sequence
20	3376.6	41.1	3482	6	AX002804	AX002804 Sequence
21	3376.6	41.1	3482	14	PERENV2	Y12239 Porcine end
22	3295.2	40.1	4918	6	AX052637	AX052637 Sequence
23	2916	35.5	3320	6	A66551	A66551 Sequence 1
24	2548.4	29.0	3025	4	AF147808	AF147808 Sus scrof
25	2416.2	31.4	15845	2	AC079471	AC079471 Sus muscu
26	2295.8	28.0	8655	6	AR136159	AR136159 Sequence
27	2295.8	28.0	8655	14	AF053745	AF053745 Sus dunali
28	2286.6	27.9	8431	4	AF151794	AF151794 Phascolar
29	2279.6	27.6	8379	14	GIU060065	U60065 Gibbon leuk
30	2261.6	27.8	214764	2	AC092712	AC092712 Mus muscu
31	2256.2	27.5	218993	2	AL592224	AL592224 Mus muscu
32	2244.4	27.3	8098	14	PGCGPE	M6927 Gibbon ape
33	2244	27.3	177788	2	AC069310	AC069310 Mus muscu
34	2244	27.3	185872	2	AC068810	AC068810 Mus muscu
35	2237.8	27.3	199883	2	AC067376	AC067376 Mus muscu
36	2235	27.2	207677	2	AL591952	AL591952 Mus muscu
37	2236	27.2	203155	2	AC079473	AC079473 Mus muscu
38	2233.2	27.2	185548	10	AC005743	AC005743 Mus muscu
39	2226.8	27.1	202333	2	AC080144	AC080144 Mus muscu
40	2221.2	27.1	217034	10	AC012147	AC012147 Mus muscu
41	2220.6	27.1	196755	2	AC093330	AC093330 Mus muscu
42	2217.4	27.0	129317	2	AC074337	AC074337 Mus muscu
43	2217.2	27.0	188460	2	AC079560	AC079560 Mus muscu
44	2216.8	27.0	188550	2	AC079547	AC079547 Mus muscu
45	2215.6	27.0	219294	2	AC015797	AC015797 Mus muscu

RESULT	1			
LOCUS	A66553	8205 bp	DNA	PAT
DEFINITION	Sequence 3 from Patent WO9740167.			
ACCESSION	A66553			
VERSION	A66553.1	GI:4538106		
KEYWORDS	.			
SOURCE	porcine endogenous retrovirus.			
ORGANISM	porcine endogenous retrovirus			
	viruses; Retroid viruses; Retroviridae; Mammalian type C			
	retroviruses; I-Mammalian type C virus group.			
REFERENCE	1 (bases 1 to 8209)			
AUTHORS	Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.			
TITLE	PORCINE RETROVIRUS			
JOURNAL	Patent: WO 9740167-A 3 30-OCT-1997;			
	Q ONE BIOTECH LTD (GB)			
FEATURES	Location/Qualifiers			
source	1..8209			
	/organism="porcine endogenous retrovirus"			
	/db_xref="taxon:61673"			

## CDS

588..2162  
/codon\_start=1  
/product="VIRAL CORE PROTEIN"  
/protein\_id="CAB39347.1"  
/db\_xref="GI:4538107"  
translation="MGQVTPPLSLTLDHWEVRSRAHNLVQYKKPKWQTFCCSEMP  
TFDYQWSEGFENSEITLAKAIIITFQPGSHPOQERYILTWDLADPWPWKWMLN  
KPKRGPRILALGSRKNHSAKEVPSSSYLPDRGADLAGTPTCSPTPLSTGCCBG  
TSAPGAVVGGPAGPAGSRGATPERTDEIALLPTKYGPPMGGOLOPQWPFSS  
ADLNMTKNHPFSEDEPQRLGLVSEIMFESHQPDCCOOLQTLFTTEERETLEA  
RKNVGADRGPTOLONEIMGFPLTRGWDYNTGDESKIKIYQALVATLRGSRP  
TNLAKREVWGPPNPPSVLELMEAFRRPTNPPDESKIKIYQALVATLRGSRP  
KKIQRIGLOAEELRDLYREAEKYTYRRETEKEQKREERERERERERKQKNT  
KIILAAVBCSSRREREDFRKIRSGPQSNLGNRLDLDQCAVCEKKGHMANRCKP  
KGNKGPVLALEEDK"  
<2163..5747  
/codon\_start=1  
/product="REVERSE TRANSCRIPTASE"  
/protein\_id="CAB39348.1"  
/db\_xref="GI:4538108"  
translation="GRGSDPLPEPRVTLKEGQVEFLVDTGAHSLVLLQPLGKLE  
KSWMGATGQROYPWTRRTVRLDVGVRVHSEFLVPECPVLLGRLLTKMAQISE  
EOGREVSNNKPTITVLLDDEYRLYSPQPDIOISWLOFROAMETAGMKIA  
KQVPOVLOLKASATPYSVROYPLSRARSGINPHVQRLIQCGILVPOSFMTPLP  
VRKGTNDYRPVDLREYNKRVODIHPVNPYNLSALPEBNWYTVLDLKAFFCL  
RLHPTSQLPFAFERDPTGRTQQLWTLRPOGFKNSPTFEDRNHLRDLANERIQHO  
VTLQYVDDLALGATKQDLLEGKALLSLDLSYASAKKQIQREYVYLSIR  
GGQWLTERRKTYVQIPATPAKQVREFGTAGFCHLPGFATLAAPLYPLTKRG  
GFSWAPHOAKAFDIAKALLSAPALAPDVTKPTLYDERKGVAGVLTQTLGPKRR  
PVATLSKLDPLVASGMPVCKATAAVALIKODKILTLGNTIVAPHALENTVRP  
DRWNTNRMTYHSLITERVTFAPRALPAPLLEDEPVTCHHOLLIEETVR  
KDLTDLPLEGEVLTWFDSSYVEEGRMAGAENVGTRIMASSLPBGSAQAEIM  
ALTOALAEAGKSNITYDSRYAFATVHGAIVKQGLTSGREIKKKEELISLE  
ALHLPKRLAIHCPGHQKADILSRGNOMARVAKOAAVNLPIETPKAPEPRRO  
YTLDEMOEIKIDOFSETPGTCYTSGLIHPKGLLEVQOIHHTLGTGTHQOL  
VRTSPYHURLPGVADSVYKHCYCOLYNNPISIPKGRILSRHSGAHMEDVTEPA  
PAKYGKYLIVPDTSGWEANPTKKEISTYVAKILIEIPRSGIPVYDSNBP  
FYAQSGLAKIIGIDMKLCARPOSSQVEAMNTIKETLTKTEGINDMALL  
PEVLPFRANTPGFGLPELLEGPPLEALFAHSADVLLSQPLFLEKALEWRO  
RAMKQLEAVSGDLOVHRFQVSDSVYRHRAGMLERMKPPLYVLLTTPPAVVE  
GIPTWIAHSHVRAPPDSGMRKTEENPLKLHLRLVYNNNSGQ"  
5620..7590  
/codon\_start=1  
/product="VIRAL ENVELOPE PROTEIN"  
/protein\_id="CAB39349.1"  
/db\_xref="GI:4538109"  
translation="MHPILSRHLPTRGEPEKRLRLPLSFASIAMFLVLTTPQASK  
RLDSSNPHRPLSLTWLIDPDTGVNSTRGVAPRGTMPELHCLRLINAVKSTP  
PMLVRSYGFCCGCTEKERYCGSGSGSPCRNSCVISNGDMKWPISLDDRKFSVN  
SGPGKKMKMLYKDKCSPSDIYLIKISTERTGYSKVDWYELGNSFLYGGAGS  
STILRLRLTETGEPVAMPDPVLAEGQPALPEPHNLPVDMTELSGFLYGGAGS  
TTLPLTNTNPNRSPGVKVGQRLPSLIGAFQALISTPDATSSCWLCSGPPYS  
GNAKERKFVYTKEHRNOCITGSRNKLITLEVGAKICSTKAPSHOHLCSYVVEA  
SENOYLAVPGNRMACNTGLTGVSVTSVNOGSKDFVMOVILPRVYHREEVLEDEY  
YRNRPKREPVSITLAVMGLGTAVGVGTALITLGPQDLKGEELHAATVEDIRA  
LKESVSNLESLSLSEVYLQNRGLDLFLRDEGGCALKECCQYVDSAIIDSM  
NKLRLKRLERRRRERADQGMFEGEMFRSMTMTLLSALNGPLVLLILTVGPCLINR  
EVAVERERSAVQIMVLRQOYGLISQGETDL"

## CDS

BASE COUNT 2168 a 2064 c 2154 g 1823 t  
ORIGIN  
Query Match 100.0%; Score 8209; DB 6; Length 8209;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 gtagttagactgtggagccagcgactgtgaataaataatcctctgtctgttcac 60  
1 GTGCTGACGACTGTGGGCCCCAGCGCTTGGAATAAATCCTTGCTGTTTCACAC 60  
QY 61 aagacgcctctcgtgagtagtgcgctcttcgagccgagacagagggat 120  
11 AAGACGCGCTTCTGCTAGTAGTATTTGGGGGTCTCCGCTCTCCGAGCCGAGCGAGGGGAT 120  
Db 61 AAGACGCGCTTCTGCTAGTAGTATTTGGGGGTCTCCGCTCTCCGAGCCGAGCGAGGGGAT 120

QY 121 tgtcttctacgaccttcaattggtgcgttgagccgaggaatccttgcagcaacctta 180  
121 TGTCTTCTTACTGCGCTTTTCATTTGGCTGTGGCCGGAAATCTCGACCACTTCA 180  
Db 121 TGTCTTCTTACTGCGCTTTTCATTTGGCTGTGGCCGGAAATCTCGACCACTTCA 180  
QY 181 caaccgagaaacgacttgagagtaaaaggaatccaccttggaagcgtgtgtgtcgcc 240  
181 CACCCGAGAACCGACTTGAGAGTAAAGAGATCCCTTTGAGAGCTGTGTGTGTGGCC 240  
Db 181 CACCCGAGAACCGACTTGAGAGTAAAGAGATCCCTTTGAGAGCTGTGTGTGTGGCC 240  
QY 241 ggcgtctctgttctgagatgtctgttctcggtatgagcgcttcggtttgacgttcc 300  
241 GGCCTCTCTGTCTGAGATGCTGTGTTCGGGTGATGCGCGCTTTCGGTTTGCACCTGCT 300  
Db 241 GGCCTCTCTGTCTGAGATGCTGTGTTCGGGTGATGCGCGCTTTCGGTTTGCACCTGCT 300  
QY 301 ctcaagacgtaagacgtgagagactgtacacagacagcgtgtcaggagatcacagcgt 360  
301 CTCAGACCGTAAAGACTGGAGGACTGTGATCAGACAGACGCTGTAGAGAGATCACAGGCTG 360  
Db 301 CTCAGACCGTAAAGACTGGAGGACTGTGATCAGACAGACGCTGTAGAGAGATCACAGGCTG 360  
QY 361 caaccctgggagacgccccggaggttgagagagacagagacgctgtgtctcact 420  
361 CAACCTGCGGAGACGCCCCGGAGGTGGGAGAGCCAGGAGCGCTGTGTCTCTACT 420  
Db 361 CAACCTGCGGAGACGCCCCGGAGGTGGGAGAGCCAGGAGCGCTGTGTCTCTACT 420  
QY 421 gtcgttgaagagacgagatctgtgttgaagcgaagcttccctccgagcgctccg 480  
421 GTCGCTGAGAGCAGAGTCTGTGTGAAGGAAAGCTTCCCTCCGCGCGCTCCG 480  
Db 421 GTCGCTGAGAGCAGAGTCTGTGTGAAGGAAAGCTTCCCTCCGCGCGCTCCG 480  
QY 481 acctcttgctctgttgagaaagcggagcgagctgcgtgtctgtatctgtgttt 540  
481 ACCTCTTGGCTCTGTGTGAGAACCGGAGCGGCTGCTGTGTGTGTGTGTGTGT 540  
Db 481 ACCTCTTGGCTCTGTGTGAGAACCGGAGCGGCTGCTGTGTGTGTGTGTGTGT 540  
QY 541 ctgtctcgtgtgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 600  
541 CTGTCTCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600  
Db 541 CTGTCTCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600  
QY 541 ctgtctcgtgtgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 660  
541 CTGTCTCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
Db 541 CTGTCTCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
QY 601 tgactcccccttaattgactcttgacatcttgacatcttgacatcttgacatct 660  
601 TGACTCCCCCTTAATTGACTCTTGACATCTTGACATCTTGACATCTTGACATCT 660  
Db 601 TGACTCCCCCTTAATTGACTCTTGACATCTTGACATCTTGACATCTTGACATCT 660  
QY 661 attgcaattcaaggttaagaaagacgttgcaagctctgtctcttgatggcga 720  
661 ATTGCAATTCAAGGTTAAGAAAGACGTTGCAAGCTCTGTCTCTTGATGGCGCA 720  
Db 661 ATTGCAATTCAAGGTTAAGAAAGACGTTGCAAGCTCTGTCTCTTGATGGCGCA 720  
QY 721 catctgaatctgagatgagcaatgagagagacatttaattctgaaatctctgtctgt 780  
721 CATTCGATCTTGATGAGCAGATGAGAGGAGACCTTAATTCTGAATTAATCTGCTGTTA 780  
Db 721 CATTCGATCTTGATGAGCAGATGAGAGGAGACCTTAATTCTGAATTAATCTGCTGTTA 780  
QY 781 aagcaatcttctgaacttgagacccggtctcattccttgataggaagcctatcctta 840  
781 AAGCAATCTTCTGACTGAGACCCGCTCTCATCTGATCAGAGAGCCCTATATCTCTTA 840  
Db 781 AAGCAATCTTCTGACTGAGACCCGCTCTCATCTGATCAGAGAGCCCTATATCTCTTA 840  
QY 841 cgttgcaagatttggagagaagatccctccgcaatgggttaaacatctggtctaaataacaa 900  
841 CGTGGCAATTTGGCAGAGAAGATCTCCGCCCATGGGTTAAACCATGGCTAAATAAACCA 900  
Db 841 CGTGGCAATTTGGCAGAGAAGATCTCCGCCCATGGGTTAAACCATGGCTAAATAAACCA 900  
QY 901 gaaagcagatcccgaaatcctgtgctcttgagagagaaacaaacactcggccgaaag 960  
901 GAAAGCAGATCCCGAATCCTGTGCTTGAGAGAAACAAACACTCTGCGCGGAAAG 960  
Db 901 GAAAGCAGATCCCGAATCCTGTGCTTGAGAGAAACAAACACTCTGCGCGGAAAG 960  
QY 961 tggagccctctcctcgatatacaccgagatcgagagcgcgacttgccggaaccc 1020  
961 TCGAGCCCTCTCTCTGATATACCCCGAGATCGAGAGCGCGGCACTTGGCGGAAACC 1020  
Db 961 TCGAGCCCTCTCTCTGATATACCCCGAGATCGAGAGCGCGGCACTTGGCGGAAACC 1020  
QY 1021 caactgttcccccacccctatcacaagaggtgtgttgagagagacctctgacctc 1080  
1021 CAACCTGTTCCCCACCCCTTATATCAGACAGAGGCTGTGTGAGGAGACCTTGCCTTC 1080  
Db 1021 CAACCTGTTCCCCACCCCTTATATCAGACAGAGGCTGTGTGAGGAGACCTTGCCTTC 1080  
QY 1081 ctgagacgtccggtgtgtgagagacgtctgcggagactcggagccgagagcgccaccc 1140  
1081 CTGAGACGTCCGGTGTGTGAGAGGACCTGCTGCGGAGCTCGAGCGGAGAGCGGCAACC 1140  
Db 1081 CTGAGACGTCCGGTGTGTGAGAGGACCTGCTGCGGAGCTCGAGCGGAGAGCGGCAACC 1140  
QY 1141 cggagcggagagagagatcgatataacgctgtgcgactaagccttccatgcag 1200  
1141 CGGAGCGGACGACGAGATCGGATATTAACGCTGCGCAGCTATGAGCCCTCCATGCGAG 1200  
Db 1141 CGGAGCGGACGACGAGATCGGATATTAACGCTGCGCAGCTATGAGCCCTCCATGCGAG 1200

QY 1201 ggggccaattgcagccctccagatctggcccttctctcgcagatctcatalatga 1260  
Db 1201 gggggcaattgcagccctccagatctggcccttctctcgcagatctcatalatga 1260  
QY 1261 aaactaaacatcccccttctcgcagatccccaagccctcaggggttgtagtccc 1320  
Db 1261 AAACAAACATCCCTCTTCGGAGATCCCAACCCCTCACGGGGTGGTAGAGCC 1320  
QY 1321 ttatgtctccacagcctcttgagatgtgtcaacagctctcagacacttca 1380  
Db 1321 TTATGTCTCTCACAGCCTCTTGAGATGTGTCAACACCTCTCTCAGACCTCTTCA 1380  
QY 1381 caaccgagagcagagagaaattctgttagagctagaaaaaatgttcctgggagcag 1440  
Db 1381 CAACCGAGAGCAGAGAGGAATTTCTTAGAGCTAGAAAAAATGTTCTGGGGCCGAGC 1440  
QY 1441 gggcagccagcagcttgcaaaatagatctgagatctggaattcccttgactcgcccggt 1500  
Db 1441 GGCAGCCAGCAGATTTGCAAAATGAGATGACATGGGATTTCCCTTACTCGCCCGGTT 1500  
QY 1501 ggaactcaacaacagcgtgaaggttagggagagcctgaaaatctatcgcagagcttggtg 1560  
Db 1501 GGAATCAACACGGCTGAAGGTAGGAGAGACTTGAATACTATCGCCAGGCTTGGTGG 1560  
QY 1561 cgggtctccggggcgccctcaagacgcccactaatttggttaaggtlaagaggtgagtc 1620  
Db 1561 CGGTTCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGGTAAAGAGGTGATGC 1620  
QY 1621 agggacccgaagaacccctccctcggtatcttcttgagaggtcctcagagcctcagcggt 1680  
Db 1621 AGGACCCGAAAGAACTCCCTCGGTATTTCTTGAGAGGCTCATAGAAAGCTTCAAGCGGT 1680  
QY 1681 tcaccccttttgatccctacccagagcccaagcctcaagtgccttgcccttgatga 1740  
Db 1681 TCACCCCTTTTGATCCTTACCTCAGAGGCCCAAGAAAGCTCATAGGCCCTTCACTTGC 1740  
QY 1741 ggcagctcggtccttgatatacagagaagaacttcagagacttgaagaggttlaacagagctg 1800  
Db 1741 GGCAGTCCGGCTCTGGATATCTAGAAAGAACTTCAGAGACTGGAAGGGTTACAGAGGCTG 1800  
QY 1801 agttacgtgatctagtgtagagagagcagaagagtgatctacagaagaggaacagaagag 1860  
Db 1801 AGTTACGTGATCTAGTAGAGAGGACAGAGAGGTGTATTACAGAAAGGAAACAGAAAGG 1860  
QY 1861 aagaagacagagagaagaagaagagagagaagaagaagaagaagcgtatagaagcgtc 1920  
Db 1861 AGAAGGAACAGAGAAAAGAAAGAGAGAGAGAAAGGAGAAACGCTGATAGAGCGGC 1920  
QY 1921 aagaagaagaatttgactaagaatcttgccgcagtggttgtagaagggaaagcagcagggaga 1980  
Db 1921 AAGAGAAAGATTTGACTAAGATCTTGGCCGAGTGGTTGAAGGGAAGGACAGCAGGAGGA 1980  
QY 1981 gagagagagatcttagaagaataatagtcagagccctagacagtcagggaaacctggcaata 2040  
Db 1981 GAGAGAGAGATTTAGGAAAAATTAGGTAGGCCCCCTAGACAGCTCAGGGAAACCTGGGCAATA 2040  
QY 2041 ggaaccacactgcagaagagcagcagctgctatgttlaaagaaaaagagcacttgagcaaga 2100  
Db 2041 GGACCCCACTGCACAAAGACAGTGTGCTGATTTGAAGAAAAAGGACACTGGGCAAGGA 2100  
QY 2101 actgccccaaagaggaacaaagagacagaggtcctagctcttagaagaagataaagatc 2160  
Db 2101 ACTGCCCAAGAAAGGAACAAAGACGGAAGGTCTAGTCTTAAGAAAGATTAAGATT 2160  
QY 2161 agggagagacggggttcggagccctcccgagcccgaggttaacttgaaggtggagggc 2220  
Db 2161 AGGGAGAACGGGTTTCGGAGCCCTCCCGACGCCAAGGTAACTTTGAAGTGGAGGGGC 2220  
QY 2221 aaacagttgagttccttggttagatccgagcagcagcaltcaagtgctgtcctaaccatag 2280  
Db 2221 AACCAAGTTGAGTCTCTGTTGATACCGGAGCGGAGCAITACAGTGTCTCAACCAATTAG 2280  
QY 2281 gaaaactaaagaaaaaactcctgggtgtagtggtgtagcagagggcaacgctagatcat 2340

Db 2281 GAAAACTAAAGAAAAAATCTCGGTGATGGGTCCACAGGGCAACGGCAGTATCCAT 2340  
QY 2341 ggaactcccgaaagaacggttagcttgagggtggagaggtlaacccactggtttcgtgtca 2400  
Db 2341 GGACTACCCGAAGAACGTTGACTTGGAGTGGAGGGGTAAACCCACTGTTCTGTGTGA 2400  
QY 2401 tcccttagtggccagtagtaccctcttagttagagactactgacccaagatggagagctcaaa 2460  
Db 2401 TCCTTAGTGCAGTAGTACCCTTCTTAGTAGAGACTTACTGACCAAGATGGAGCTCAAA 2460  
QY 2461 ttctcttgaacaagaagaaccagaagtgctgttgaaataaacaaccactactgtgttga 2520  
Db 2461 TTTCTTTTGAACAAGAGAACCCAGAGTGTCTGTGAATTAACAAACCCATCACTGTGTGA 2520  
QY 2521 cccctcaattagatgtagatatacgactatcttccccaagttaaagcttgatcaagata 2580  
Db 2521 CCTTCAATTTAGATGATGATATGACATATTTCTCCCAAGTAAAGCCTGATCAAGATA 2580  
QY 2581 tacagtcclggttgagagcagtttccccaagcctgggcaagaacccgaggtatgggttgg 2640  
Db 2581 TACAGTCTGTGTGAGACACTTTCCTCAAGCCTGGGCAACACCGAGGGATGGTTGG 2640  
QY 2641 caaagcaagtlcccccacaggttatcactgaagggccagtgctacacagatcaagtica 2700  
Db 2641 CAAGCAAGTTCCTCCACAGAGGTATTCAACTGAAGGCCAGTGTACTACACAGTATCAGTCA 2700  
QY 2701 gacagtagcccttgtagtagagaggtcgcgaagaggaatttggccgcatgttcaagaattaa 2760  
Db 2701 GACAGTAGCCCTTGATGATGAGAGGCTGAGAAAGAAATTTGGCCGCAATGTTCAAAAGTTAA 2760  
QY 2761 tccaacagggcatccctagttccctgtctcaatcccttggaactactccctgctacaggtta 2820  
Db 2761 TCCAACAGGGCATCCTAGTTCCTGTCTCAATCCCTTGGAAATPACTCCCTGCTTACCGGTTA 2820  
QY 2821 ggaagccttggagccaatgtatlcagaccagtagcaagacttgaagaggtcctaaataaagg 2880  
Db 2821 GGAAGCCTGGAGCAACATGATTTATCGACAGTACAGAGACTTGAAGAGGTGCATTAAGAGG 2880  
QY 2881 tgcaggaactataccccaagcgtccgaacccctatacctcttgagcgccctccgctg 2940  
Db 2881 TGCAGGACATACACCCCAACGGTCCCAACCTTAAACCTTGAGACGGCCCTCCGCTCG 2940  
QY 2941 aacggaactgtgaacagatattggaacttaaaagatcccttctcgcgtgagattacacc 3000  
Db 2941 AACGGAACCTGTACACAGATTTGACTTAAAGATGCTCTTCTTGCTCGAATTTACACC 3000  
QY 3001 ccaactagccaacacttcttgcttcgaaatgagagatcccaagtagcgggaagaacccgggc 3060  
Db 3001 CCACCTAGCCAACACACTTTTTCCTTGCATGAGATGAGATCCAGGTACGGGAACACGGGSC 3060  
QY 3061 agtcaacttggaccgcacttgcgcccaagggttcaagaactcccgacacatctttagcgaag 3120  
Db 3061 AGCTACCTTGAGACCCCACTGCCCCCAAGGGGTTACAGAACTCCCGACCATCTTTGACGAAG 3120  
QY 3121 cccctacagaagagacttgcgcaacttcgaagtagtcaaacaccttaagttgacccctccagat 3180  
Db 3121 CCTTACACAGGACCTGGCCAACTTGAATGATCCAAACCTTAGGATGACCTCTCCAGT 3180  
QY 3181 acgtgatalacacttgccttcggcgaggagccaacaagaagctttagaaggtlacgaag 3240  
Db 3181 ACGTGATGACCTGCTTCGCGGGAGGACCAACAACAGGACCTGTTAGAAAGGTACGAAGG 3240  
QY 3241 caactactgtgaattgttgaactagcttagacagacgtctgtctaagaagccagattc 3300  
Db 3241 CACTACTGCTTGGAATTTGCTGACTTGGCTTACAGAGCCCTCTCTAAGAAAGGCCAGATT 3300  
QY 3301 ggaaggaaggtlaacataacttggggtlaagtttcggggcgaggagagatgtgcgaagg 3360  
Db 3301 GCAAGGAGAGGTTAACTATCTTGGGGTACAGTTTCCGGGGCCGGACGAGTGTGACGG 3360  
QY 3361 aggcaggaagaagaactgttagtcagatccggccccaacccagcgaacagctgagag 3420

[illegible]





```

Db 7741 GGTGAAGTACACACTTCTATTGTTCCAGGCGCTGCTATCTGGCGCTAAGTAAGTAAC 7800
Qy 7801 aggaatgagtgactaatcgtctatctggtatctgtlaaactgactgagccatagaag 7860
Db 7801 AGGAATGAGTGTGACTAATCGCTAATCTGATCTGTAAACTGACTGGACCATAGAAG 7860
Qy 7861 aattgatacaattgacagccctagtgtacctatctcaactgcaatctgtacctgccc 7920
Db 7861 AATTGATTACACATGTGACACCCCTAGTACCTTCACATGCAATCTGTCACTCTGCC 7920
Qy 7921 agagccacagatcgagactccgagactatlttaaatgattgtgtccacgagagcgc 7980
Db 7921 AGAGCCACAGCATGACATCGGACTCCGAGACTATTTAAATATGTTGTCACAGGAGCC 7980
Qy 7981 gggctctgataatttaaatgaltgtgtccatgagagcgcggtctcgatalttaaat 8040
Db 7981 GGGCTCTGATATTTTAAATGATGTGATCATGTGAGCGCGGCTCTGCAATATTTTAAAT 8040
Qy 8041 gattgtgtgtgacgacagcttgtgtgtgaaccccatgaagctgtccgacttcgca 8100
Db 8041 GATTGTGTGTGACGACAGGCTTGTGTGAACCCCATGAAGCTGTCCGATTCGCGA 8100
Qy 8101 ctgagagccgagtcctctccctcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8160
Db 8101 CTGAGGCGCGGACGACTCTCTACCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8160
Qy 8161 aataaatacctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8209
Db 8161 AATAAATACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8209

RESULT 2
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus.
VIRUSES; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
FEATURES
source Q ONE BIOTEC LTD (GB)
location/Qualifiers
1..8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

```

```

Query Match 99.5%; Score 8171.4; DB 6; Length 8196;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 8191; Conservative 4; Mismatches 0; Indels 2; Gaps 2;

```

```

Db 181 GACTTGAGAGTAAAGGATCCCTTTGAGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Qy 253 ctgagtgatctgttctcgttgatgcgagcttccgtgttgcagctgtctctcagaccgttaa 312
Db 241 CTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy 313 ggaactggaagactgtgtatcagcagaagctgtctagaagatcacaagctgtccacctgggg 372
Db 301 GGAAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Qy 373 agagcccggaaggttgagagagccagagacccctgtgtgtgtgtgtgtgtgtgtgtgtgt 432
Db 361 AGGCCCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Qy 433 accgagttctgttctgaaagcgaagcttccctcccgagcggtccgactcttggcct 492
Db 421 ACCGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 493 gcttctggaagagcgagagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 552
Db 481 GCTTGTGGAAGACGGGAGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 553 tcttctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 612
Db 541 TCTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 613 ttagttgactctcagacatgtgactgaagttagatccagagctcactaattgtcagttc 672
Db 601 TTAGTTGACTCTCAGACATGTGACTGAAGTTAGATCCAGGGCTCATATTTTGTACATTC 660
Qy 673 aggttaagaagagagactctgtgagactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 732
Db 661 AGGTTAAGAAGGAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy 733 gatggccatcagaagagagacatttaattcgaatatactgtgtgtgtgtgtgtgtgtgtgt 792
Db 721 GATGGCATCAAGAGGAGACCTTATTTCTGAATATTCCTGCTGTGTGTGTGTGTGTGTGTGT 780
Qy 793 ttcagactgagccgagctctcactatcctgatacagagacccatacttaagctgtgcaagat 852
Db 781 TTCAGACTGAGACCCGCTCTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 853 tggcagaagatcctcgcgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 912
Db 841 TGGCAGAGATCTCTCCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy 913 cccgaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 901 CCCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy 973 cctgtatctaccccgagatcagagagccgcgcgactgtgtgtgtgtgtgtgtgtgtgtgt 1032
Db 961 CCTGTATCTACCCCGAATGTGAGAGCGCGGACTTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Qy 1033 ccaacccttaccagacagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1092
Db 1021 CCACCCCTTATCCAGCAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy 1093 tggtagagagactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1152
Db 1081 TGGTGAAGGAGACTCTCCGAGACTGTGAGACCGGAGAGGGCCACCCCGGAGGGAGAG 1140
Qy 1153 acgagatcgagatattacagctgtgcacatgtgacctccacatgcaaggggcaattgc 1212
Db 1141 ACGAGATGCGATATTATACCGCTGTGCGCACCTTATGTGCGCCATGCCACAGGGGCAATTGC 1200
Qy 1213 agccctcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1272
Db 1201 ACCCCCTCAGATTGTGCGCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 1273 ccccttctcgaagatcccaacgctcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1332
Db 1261 CCCCTTCTCGAGAGATCCCAAGCCCTCAGAGGGGTTGTGTGTGTGTGTGTGTGTGTGTGT 1320

```

QY	1333	acgcgcctacttggatgatgtgtcaacagcctgcctgcagacactcttcaaacgcgaagac	1392
Db	1331	ACCGACCTACTTGGGATGATGTTCACACACTCTGTAGCACCTTTCACCAACCGAGGAGC	1380
QY	1333	gagagaaattctgttagaggtctagaanaaatggttccttggggcgagacggcgaccacgc	1452
Db	1381	GAGGAGAAATTCTGTTAGAGCTAGAGAAAAAATGTTCTGGGGCGCAGCGGGGACCCACGC	1440
QY	1453	agttgcgaatagatgtgaatbgtgatcttcccttgactgcgcgcgggttggagctacaaca	1512
Db	1441	AGTTGCAAAATGACATTGACATGGGATTTCCCTTGACTGCCCGGCTTGGGACTACACA	1500
QY	1513	cgctctgaagtlaaggagaagcttgaaaatcatatccagctctgtgtgcgggtctcggg	1572
Db	1501	CGGCTGAAGGAGGAGAGAGCTTGAATAATCATCGCAGAGCTGTGTGGCGGGCTTCCGGG	1560
QY	1573	gcgcctcaagaacggccacctaattgtgtcctaaggtaagagagtgaatgaaggagaccgaa	1632
Db	1561	GGCGCTCAAGACGGCCCACTAATTTTGGCTTAAGAGAGGTGATGACGAGGACCGAAGC	1620
QY	1633	aacctccctgcgtattctcttgaagagcctcatbgaagcccttgaagcggttcaaccctttg	1692
Db	1621	AACCTCCCTCGGATTTTCTTGAGAGCCTATATGAAACCTTCAGCGGGGTTCAACCCCTTTG	1680
QY	1693	atccctacctcagaagcccacagaagaacctcagcttgcccttgctctcatltggcagctcgc	1752
Db	1681	ATCCTACCTCAGAGGCCCAAGAACCTCATGTGGCCCTGGCCTTCATGTGGCAGTGGGCTC	1740
QY	1733	tggatatacagaagaaccttcagagaccttgaagaggtttacagaagcttgaatgaatgac	1812
Db	1741	TGGATATCAGGAAGAAACTTCACAGAGCTGTGAAGGGTTACAGAGGCTCAGTTACGTGATC	1800
QY	1813	taftgagagagggcgagaaagttgtattcaagaaggagacaaagaagagaagaagaacga	1872
Db	1801	TAGTGAAGAGGCGAGAAAGGTTATTACAGAGGGAGACAAAGAGAGAGAGAGACAGA	1860
QY	1873	gaagaagaaggaagaagaagaagaaggagagaagaacgtatataagacgcgaagaagaat	1932
Db	1861	GAAGAAGAAAGGAGAGAGAAAGGAAGGAAGGAGGAAGAGCTGATAGACGGCAGAGAAATTT	1920
QY	1933	tgactaatatcttgcgcgcagtggttctgaaggagaagacagcagagagaagagagagatt	1992
Db	1921	TGACTAAGATCTTGGCCGCGAGTGGTTGAAGGGAGAGCGACGGAGAGAGAGATTT	1980
QY	1993	ttaggaanaattagtgagcgcccttagacagttcagaagaaaccttggcgaatbgaagcccaatc	2052
Db	1981	TTAGGAAATTAAGGTCTAGCGCCCTTAGACAGTCAGGCAACTGGGCAATTAAGACCCCATCG	2040
QY	2053	acaagagacagtgctgtcatctttaaagaaanaagacaactgggcgaagacttcccccaaga	2112
Db	2041	ACAAGGACCACTGTGCTATTGTTAAGSAAAMAAAGACACTGGGCAAGGAACCTGCCCAAGA	2100
QY	2113	agggaaacaagaagcgcgaaggtcctagctctataagaagaataaagattagggagaaggg	2172
Db	2101	AGGGAAACAAGGAGACCCAGA-GTCTAGTCTCTAAGAAAGATTAAGGGAGAC-CG	2158
QY	2173	gttgcgacccctcccggaacccaaggtiaactttgaagtggaaggggcaaaccaagtgtgt	2232
Db	2159	GTTGGGACCCCTCCCGAGCCCAAGGGTTAAGGAGGTGAAGGGGAGCGGCAACCATGTTGAT	2218
QY	2233	tcctgtgtgataccggagcggaagatctcaatgtctgtataacaacatlaagaaacaaag	2292
Db	2219	TCCTGTGTGATACCGGAGCGGAGCATTCATGCTGTACACCATTTAGSAAAACTAANAAG	2278
QY	2293	aaaaaaatctctgggtgaatgtgttccaagggcaacgycagtatccatgactaccgaa	2352
Db	2279	AAAAAAATCTCTGGGTATGGGTGCCACAGCGCAACGGCAGTATCCATTGACTACCCGAA	2338
QY	2353	gaacgtgtgacttggaggttggagacgggttaaccacactcgtttctgtfcaatccctagtgcc	2412
Db	2339	GAACGTTGACTTGGAGGTGGGAGCGGGTTAACCCACTGTTTCTGGTCATCCTTGATGCC	2398

QY	2413	cagtaacccctctaagtagaactactctyacaagaatggagctccaaattcttttgaac	2472
Db	2399	CAGTACCCCTTCTAGAGTAGAGACTTCTCACCAAGATGGAGCTCAAAATTTCTTTGTAAC	2458
QY	2473	aaggaagaaccaagaagctcgtcgttaataaacaacccaatacgtctgtttgaacctccaatag	2532
Db	2459	AAGGAAGAACCAGAAGTCTCTGTAATAAACAACCAATCACTGCTGTGACCTCCAAATTAG	2518
QY	2533	atgtatgaatcgaactatcttccccaagttaaagccgtatcaaatatlaacgtccgt	2592
Db	2519	ATGTATGAATATTCAGACTATATTCGCCCAAGTAAAGCTGATCAAGATATACAGTCTGGT	2578
QY	2593	tyageacgattccccaagccttggcagaaaccgcagagatgggtttgcaaaagcttc	2652
Db	2579	TGGAGCAGTTTCCCAAGCCTGTGGGCAAGAAACCGCAGGAGATGGGTTTGGCAAGCAAGTTC	2638
QY	2653	ccccacaggttatctaacttgaagccaagctctaaacccaagtaicagttaagaagtaacct	2712
Db	2639	CCCCACAGTTATTTCACTTAAGCCCAAGGCTACACAGTATCACTCAAGAGTACCCCT	2698
QY	2713	tgagtagaagagctcgcgaagaagatttggccgacgttcaaaagatlaataccaagaagca	2772
Db	2699	TGAGTAGAGAGAGCCTCGAGAGGAATTTGGCCGCAATGTTCAAAAGATTAATTCACAGGGCA	2758
QY	2773	lccatgtctcgtlccaatcccccttggataataccctctgctacoggttaggaagccttgg	2832
Db	2759	TCTGATGTTCTGTTCCATCCCTCTTGGAAATACCTCCCTGCTACCGGTTAGGAAGCCTGGGA	2818
QY	2833	ccaatgattatcgacaagctacagacttgaagagtgtaataaagggtlgaagacatac	2892
Db	2819	CCAAATGATTATTCAGACAGTACAGAGCTTGAAGAGAGTCAATTAAGGGTGTGAGACATAC	2878
QY	2893	acccaagctccccaacccctataaacctcttagacgccccctccgcgcttlaacggaactcgt	2952
Db	2879	ACCCACAGGTCGCCAACCTTATTAACCTCTTAGAGCCCTCTCCGGCTTAACGGAACCTGGT	2938
QY	2953	acacagatatttgaacttlaaagaatgccttctcctcgttagatlaacccaactagcaac	3012
Db	2939	ACACAGTATTGAGACTTTAAAGATGCGCTTCTTGCTCGTAGATTACACCCCACTAGCCAAC	2998
QY	3013	caacttttgccttcgaaatlgagaagatccaaggttagcgggaagacccggagactaacctcgtga	3072
Db	2999	CACTTTGTGCTTTGGAATGGAAGATCCAGATCCGGAAGAACCCGGAGCTCATCTGGGA	3058
QY	3073	cccgcagcccccaagaggttlaagaactccccgcacactcttgaacgaagcccttacaagag	3132
Db	3059	CCCGACTCTCCCAAGGGTTCAAGAACTCCCCGACCATCTTTTACGACAAACCTCTACCAAGGG	3118
QY	3133	accttgcacaacttcaagatccacaacaccttaagtgaacctctccagtaagltgagacc	3192
Db	3119	ACCTGCGCAACTTTCAGATTCACAAACCCCTCAGTGACCTCTCCAGTACGTGATGATCC	3178
QY	3193	tgcttcttggcgggagccacaacaacagactcgtctagaagatlaaggaagcactactcgtcgtg	3252
Db	3179	TGCTTCTTGCGGGAGCCACAACAACAGAGACTGCTTACAAGATACGAAGAGCACTAGCTGGG	3238
QY	3253	aattcttgcaccttagctctacaagacccctcgtctlaagaagaagccagatlttgcagagagag	3312
Db	3239	AATTGTCTGACCTAGAGCTTACAGAGACCTCTGCTTAAGAAGGCCACGAATTTGCAAGAGAGAGG	3298
QY	3313	taacatacttggggatagaaglttgcgggcggcagcgtatgcttgcacggaggaacggaaga	3372
Db	3299	TAACTACTTGGGGTACAGTTTGGGGGGCGGACAGATGCTGTACGAGAGGACAGGAAGA	3358
QY	3373	aaactgtatgtccagatataccggcccccaacacagccaaacaadltgagagatlttttgggga	3432
Db	3359	AAACGTATGTCCAAATATCCGGCCCCAACACACACCAAAATGTGAGAGATTTTGGGGA	3418
QY	3433	caagctggatttgcagaagcttggatccccgggggttgcgaccttgaagaacccaacttaac	3492
Db	3419	CAGCTGGATTTTGCAGAGACTGTGGATCCCGGGGTTTTCGACACTTACAGACCCCACTATACC	3478
QY	3493	gccttaaccaagaagaaggggagatctccttgggctccttgaagaccagaaggaatttga	3552

Db 3479 GCGTACCAAGAAAAAGGGGATTCTCTGGGCTCCGACACACAGAGCATTTGATG 3538  
OY 3553 ctatcaaaaagccctgtctgagcgaaactgtctctgcccctccctgaacttaacacct 3612  
Db 3559 CTATCAAAAAGGCGCTGTGAGCGCAGCTGCTTGCCCTCCCTGAGCTTAACCTT 3598  
OY 3613 ttacccttattgtgatacgtaagagagtagccgagagatttlaaccnaaccctag 3672  
Db 3599 TTACCCCTTATGTGGATGAGGAGTAAAGGAGTAGCCGAGGAGTTTAAACCAACCGTAG 3658  
OY 3673 gacatagagagagagactgtctgcttaacctgtcaagaagactgtatccgttagcaagtgt 3732  
Db 3659 GACCATGAGAGAGAGACTGTCTTCCCTCACTGCTAAAGAGCTTGAATCCCTGAGTGT 3718  
OY 3733 ggcacctatgtctgaagactctgcaagctgtgcccatactgtgtaagaagagcttacaat 3792  
Db 3719 GGCCCGTATGTCTGAGAGCTTCCGACGCTGTGGCATACTGTGTAAGAGAGCGCTACAAAT 3778  
OY 3793 tgaacttggagacagaataaactgtlaagcccccaatgcatltagagaacatcgtlccgc 3852  
Db 3779 TGACTTTGGACAGAAATTAATTAATAGCCCCCATGATGAGAGACATGCTTGGCC 3838  
OY 3853 agccccagacagatgtatgacaaagccgcatgacccaatacaaaagcctgtctca 3912  
Db 3839 AGCCCCAGACCGATGATGATGACCAAGCCCGCATGACCCACTATCAAAAGCCTGCTTCA 3898  
OY 3913 cagaagagtgtaactctgctccacagccgctcaacccctgcaacccctctgctgaag 3972  
Db 3899 CAGAGAGGCTACCTTGGCTCCACAGCCGCTCAACCTGCGACCTCTTGTGCTTAAG 3958  
OY 3973 agactgtatgaacagtgactcaatgatactcaactatgatatgaagagagctggtlcc 4032  
Db 3959 AGACTGTATGAACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 4018  
OY 4033 gcaagagccttaccagacataccctgtctgagagagagtgctaaacctgttcaactgaagaa 4092  
Db 4019 GCAAGGACCTTACAGACATCCGCTGACTGAGAGAGTGTACCTGCTTCACTGACGAGAA 4078  
OY 4093 gcaactatgtgtgtaagagtaagagatgtgtgagcgagtggtgagagagagagagag 4152  
Db 4079 GCACCTATGTGTGTAAGAGTAAGAGATGCTGCTGGGCGGAGTGTGTGAGAGAGAGAGAG 4138  
OY 4153 cgaatcgtgagcagcagctgcccgaagaaacttaagcgaacaaagctgagctcaatgccc 4212  
Db 4139 CGATCTGGGCGCAGCAGCTGCGGGAAGAACTTACAGGCAAAAGGCTGACCTCATGGCC 4198  
OY 4213 tcaagcagacttgcgtgctgagcgaagagaaatccataaacaattatacgagacagagt 4272  
Db 4199 TCACGCAAGCTTGGCTGGCTGGCGGAAGAAATCCATTAACATTTATACGAGACAGAGT 4258  
OY 4273 atgacctgtgactgacagatcaacgagagcattcaataaagagaggttgcctacct 4332  
Db 4259 ATGCTTTGGAGCTGCAGACAGTACAGGCGCATCTTAACAAAGAGGCTTCTTAACCT 4318  
OY 4333 cagcagagagagaaataaagaacaaagagaaatctcaagcctattagaaccccttcaat 4392  
Db 4319 CAGCAGGAGGAGAAATTAAGAAACAAAGAGAAATTCATACCTATTAGAAAGCCTTACAT 4378  
OY 4393 tgcacaaaagagtagctattatactgtctctgagacatcaagaagccaaagatcctaat 4452  
Db 4379 TGCACAAAAGGCTAGCTATTATACCTGCTGACATCAAGAAAGGCAAGATCATATAT 4438  
OY 4453 ctagaagagacagatgctgacccgggttgcacagcagcagccagcagctgttaaccttc 4512  
Db 4439 CTAGAGGAGACAGATGCTGACCGGGTTCACAGAGAGGAGCGCCAGGCTTTAAACCTTC 4498  
OY 4513 tgcctataatagaacagcccaagcccaagacccagacagacaglaaaccttagaagact 4572  
Db 4499 TGCCTAATATAGAAAGCCCAAGCCCAAGACCCAGAGAGAGTACACCTTGAAGAGT 4558  
OY 4573 ggcagaagataaaaagatagacagattctctgaagctccgagagagagcctgtataact 4632  
Db 4559 GGCAGAGATAAAAAGATAGACCACTGCTGAGACTCCGAGAGGAGACTGCTATACCT 4618  
OY 4633 catatggagagaaatcctgcccccaagaagaggttaataatgtccaaacagatataac 4692  
Db 4619 CATATGGAGAGAAATCTGCCCCACAAAGAGGTTAATAATGTCCAAACAGATACATC 4678  
OY 4693 gtctaaacccactagaagactaaacacactgcaagagttgttcaagaatcccttcatgt 4752  
Db 4679 GTCTAACCCACCTAGAGACTAAACACCTGACAGAGTGTGTCAGAACATCCCTTATCATG 4738  
OY 4753 tcttgaggtacccaagagatggtctgactggtgtgtcaacaactgtgttccctgagctgtg 4812  
Db 4739 TTCTGAGGCTAACCCAGAGAGTGTGCTGACTCGGTGCTCAAACTTGTGTGCTCCAGCTGG 4798  
OY 4813 ttaatgctaatcccttcagaataaccccaagaaagagactaaaggagagagagagagag 4872  
Db 4799 TTATGCTATCTCTCCAGATACCTCCAGAGAAAGAGACTTAAGGAGAACCCACAGGGG 4858  
OY 4873 ctactgggaagtgagacttcaactgagtgaaagccgctaaataagcgaacaaatactat 4932  
Db 4859 CTCACGAGAGAGTGCAGCTTCACTGAGGTAAACCCGCTAAATACGAAACAAATATCTAT 4918  
OY 4933 tggcttctgtagaacacttccaagatggtgtagaggtctatccctactaagaagagact 4992  
Db 4919 TGGTTTTGTAGACACTTTTCAAGATGGGTAGAGGCTTATCTTAAGAAAGAGACTT 4978  
OY 4993 caacgctgtgtgtaagaataaactgagagaaattttccaagatttgaataactaagg 5052  
Db 4979 CAACCTGCTGCTTAAGAAATACCTGAGAGAAATTTTCAAGATTTGGAATACCTAAGG 5038  
OY 5053 taataggtcaagcaatgttccagcttctgttcccaagtgaaagagagagagagagag 5112  
Db 5039 TATATAGGTCAGACATGCTCCAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 5098  
OY 5113 tattggagattgattggaactgcatlgtcaatacagaccccaagctcagagagtag 5172  
Db 5099 TATTGGGATTTATTTGGAATGCAATTTGCTATACAGACCCCAAGAGCTCAGAGACGTAG 5158  
OY 5173 agagagtagaatagaacattaaagagacccctaccaaattgaaccaagaagctgcatata 5232  
Db 5159 AGAGGATGATGAAACCATTAAGAGACCTTAAACAAATGACCAAGAGATGGCATTA 5218  
OY 5233 atgattgtagtgcctccctgcttctgtcttctttagagtgagagacccctgagacgt 5292  
Db 5219 ATGATTGATGAGCTGCTCCGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 5278  
OY 5293 ttggctgagcccccatgaaatgtctcctagcgggagaccccccgttggcagaatgtcct 5352  
Db 5279 TTGGGCTGACCCCTATGAAATGCTTACGGGGAGACCCCTGCTGGCAGAAATTTGCT 5338  
OY 5353 ttgcacatagtgctgattgtgctgtctccagccttctcctagagctcaagcgtctcg 5412  
Db 5339 TTGCACATAGTCTATGCTGCTGCTTCCACACCTTTGCTTCTAGGCTCAAGGGGCTCG 5398  
OY 5413 agtgggtgagcagcagcgttgaagcagctcccgagagagcctactcaagagagagactg 5472  
Db 5399 AGTGGGTGAGCAGGAGGAGGTGGAAGAGCTCCGGGAGAGCTACACAGAGAGAGACTTGC 5458  
OY 5473 aagttccacatcgtctcaagttgagatcaagttatgttaagcagccagcgtgagaa 5532  
Db 5459 AAGTTCCACATCGCTTCCAAAGTTGAGATTCAGTATGTTAGAGCGCCACCGTCAGGAA 5518  
OY 5533 acctggaagctggtggaagagagacttctcctgaactttgaaccaaccaagcgtctga 5592  
Db 5519 ACCTGAGAGCTGAGGAG 5578  
OY 5593 aagtcgaagaaatcccccactgagatcaatcccaagcgttgaagcagcagcagcagcagc 5652  
Db 5579 AAGTGGAAGAAATCCCCACGTCGATCCATGATCCAGCTTAACAGGCGGCACTCTCCG 5638  
OY 5653 actcgggttgaagagcgaagaaagagactgaatcccttaagcttcgctcactcagctgtg 5712  
Db 5639 ACTCGGGGTGAG 5698

QY 5713 ttccctactctaaacaaatacctcccaagccagtaglaaaacgccttatagacagctcgac 5772  
|||||  
Db 5699 TTCCTTACTCTAACAATACTCCCGAGCCAGTAGTAACGCCCTATAGACAGCTCGAAC 5758  
5773 ccccaatagaccttatcccttaacctggtatattagacctgtagaggtgtagctta 5832  
|||||  
Db 5759 CCCCATAGACCTTTATCCCTTACCTGCTGATATTGACCCCTGATACCGGAGTCACTGTA 5818  
5833 aatacactcgaggtgtgctcctctagagcaacctctgagccttgaaactcattctcctc 5892  
|||||  
Db 5819 AATAGACTCTCGAGGTGCTCTCTAGAGGCACTGCTGCTGAACTGCATTTCTGCTC 5878  
5893 cgattgataaaccgcgcgtgtlaaaagcaacctcccaacctagtcgtagtattaggttc 5952  
|||||  
Db 5879 CCATTGATTAACCCCGCTGTAAAGCACACCTCCCAACCTAGTCCGAGTATAGGCTTC 5938  
5953 tattgtcccaagcaacagagaagaatactgtgtgggtttctggggaaactcctctgt 6012  
|||||  
Db 5939 TATTGCTGCCAGGACACAGAAAGAAATACCTGGGGGCTTCTGGGGAAATCCTTCTGT 5998  
6013 agagatagagctgtgctaacctcccaacagtagagactggaatgtgcgactctctccag 6072  
|||||  
Db 5999 AGAGATGAGAGCTGCTGCTCACTCCACGATGAGACTGGAATGSCCATCTCTCTCAG 6058  
6073 gaccgggtlaaaattctccttctgtcaattcccgccgggcaagtagaanaatgtagaacta 6132  
|||||  
Db 6059 GACCGGGTAAATTTCTCTTCTTGTCAATTCGCGCGGCAATGACAAATGATGAACCTA 6118  
6133 tataaagataaagctgtcccccatacagactagattatctaagaataagttcaactaa 6192  
|||||  
Db 6119 TATAAAGATTAAGAGCTGCTCCCATCAGACTTATCTTAAGATTAAGTTTCTCTGTA 6178  
6193 aagaaacagagaataatatacaaaatgtaaaatgtagaactgtgggaataagatttcta 6252  
|||||  
Db 6179 AGGAAACAGAGAAATATTTCAAAAGTGAATGATGATGAGACTGGGGAATGTTTATA 6238  
6253 ttatatgagggagagagaggtctcaacttaaccatcgctccttagatagagaagggaca 6312  
|||||  
Db 6239 TTATATGGCGGGGAGCAAGGCTCACCTTAACTTCGCTTACGATTAAGAGACGGGACA 6298  
6313 gaaacccctgtgagcaatgtagaccgataaagtactgagtggaagaagggcccccgcctg 6372  
6299 GAAACCCCTGTGGCAATGGACCCGATTAAGTACTGCTGAACAGGGGCCCCCGGCCCTG 6358  
6373 gaggcaacagataactgtgcggtgtcccaatlaaacctgcgtgcgactgacataaacaag 6432  
|||||  
Db 6359 GAGCCACCGCATTAATCTGCGGTGCGCCCAATTAACCTGCTGCGGCTGACATTAACACAG 6418  
6433 ccgcctagcaacagtaaccactgtatctccttaaccaacagcctagaanaactcccaagt 6492  
|||||  
Db 6419 CCGCTTAGCAACAGTACCACTGATGATTGATCCACCAACAGCGCTGAAACACTCCCAAGT 6478  
6493 gtcccggttaagaagaagaagagactcttcagctcatcacaggaagcttcccaagcactc 6552  
|||||  
Db 6479 GTTTCCTTTAAAGACAGAGACGCTTCATCTCATCAAGGACCTTTCCAAAGCAATC 6538  
6553 aactccacagacctgatgtcaacttctctgtgtgcttgcctcctcctcaagggcctcct 6612  
|||||  
Db 6539 AACTCCACCGACCTGATGCTACTTCTCTGTGTGCTTGTCTATCTCCAGGGCTCTCT 6598  
6613 tattatagaggagtaggctaaagaagaanaatlcactgtgacaaagaagatagaatacaa 6672  
|||||  
Db 6599 TATTATAGGGGATGGCTAAAGAAAGAAATTCATATGTCACCAAAAGAGCATAGAAATCAA 6658  
6673 tgtacataggggttcccgaaataagcttacccctacaggaagttccgggaagggagacatc 6732  
6659 TGTACATGGGGGTCCGAAATTAAGCTTACCTCATCTGAAGTTTCCGGGAAGGGGACATGC 6718  
6733 ataggaanaagctcccccatacccaacaacacttgcctatatagtagtgtgttctagaagc 6792  
|||||  
Db 6719 ATAGGAAAGCTCCCATCCCAACCAACACTTTGCTATATGTAATGTAATGAGCAG 6778  
|||||

---

QY 6793 gccctagaanaatcagtaattagtaacctgtgtataacaaggtgtgtggcatgcaatactgg 6852  
|||||  
Db 6779 GCGTCAGAAATATCAGTATTTAGTACCTGGTTATACAGAGTGGGGCATATGCAATACTGG 6838  
6853 tlaacccctgtgtttccacctagcttccaaactccaaagattctgtgtcatgttc 6912  
|||||  
Db 6839 TTAAACCCCGTGTTCACCTCAGTCTTCAACCAATCCAAAGATTCTGTGTCATGCTC 6898  
6913 caaatgtccccagagtgtactacaactccgaggaagtggttctgtatgatatgtactat 6972  
|||||  
Db 6899 CAAATCTCTCCCGAGGTACTACATCTGAGGAATGGTCTTGATGGAATATGACTAT 6958  
6973 cggfataaccgaccaaaagaagaaccgtaacctcccttaacctgaactgaatgctcgaata 7032  
|||||  
Db 6959 CGGTATACCGACCAAAAGAAAGAACCCGTATCCCTTACCTAGCTGTAATGCTCGATTA 7018  
7033 gggagcgccgtgtgagtagaacaaggagacagctgtccctgaatacaagaacacagcagta 7092  
|||||  
Db 7019 GGGAGCGCCGTGGCGTAGGAACAGGAGACTGCTGATACAGAGACCAAGACAGACTA 7078  
7093 gagaagaagactgtgtgagctacatgtgcgcacatgacagaagaatctccagagccttaagag 7152  
7079 GAGAAAGGACTTGGTGAAGCTACATGCGGCATGACAGAAATCTCCAGCTTAAAGGAG 7138  
7153 tctgttagcaacctagagaagatccctgaacttctgtcctgaatggtgtctacagaacgg 7212  
|||||  
Db 7139 TCTGTAGCAACTAGAGAGTCCCTGACTTCTTGTCTGAAGGCTTCTACAGAACCGG 7198  
7213 agggagatagatctgtcgttctlaagaagaagtgtgttattgtgacagccttaaaagaaga 7272  
|||||  
Db 7199 AGGGATTAAGATCTGCTGTTCTAAAGAAAGTGGTTATGTGACAGCTTAAAGAGAA 7258  
7273 tgtgtctctatgtatagatcaactcaagggccaatgaagactccatgaacaagccttagaana 7332  
7259 TGTGCTTCTATGTAGATCACTCAGAGCCATGACAGAACTCCAGAAACAGCTTAGAA 7318  
7333 aagttagaagagcgtcagaaaggaagaagcctacagaaggtgtgttgaagaatgttc 7392  
|||||  
Db 7319 AAGTTAGAGAGCGCTCGAAGGAAAGAGGCTGACCAAGGGGTGTTGAAGATGGTTC 7378  
7393 aacagctccttctgtagtaccacacctgtctctcctcctgacggggcccttagtactcgt 7452  
7379 AACAGTCTCTTGGATGACCACTGCTGCTGCTGAGAGGGGCCCTAGTACTCTG 7438  
7453 ctccgttacttaacagttgggctgtcttaataatagttgtgtgcttgttagaaga 7512  
7439 CTCTGTTACTTACAGTTGGCGCTTGTCTTAATATAGTTGTTGCTTGTAGAGAA 7498  
7513 cgaagtgtgtcaggtccagatcatatgtacttaggcaacagtagtaacagcctctgagcaca 7572  
7499 CGAGTGAAGTGCATGCCATATGATGTTAGGCAACAGTACCAAGGCTTCTGAGCAA 7558  
7573 ggaagaactgacactlaagccttcccaagttcctaagaatlaagaactlaacaagaagaag 7632  
7559 GAGAAACTGACCTCTAGCTTCCCAAGTTCAAGATTAAGACTATTAACAAGACAAGAG 7618  
7633 tgggaatgaaagaatgaaatgcaactaacctcccaagaacccagaagaatlaataaa 7692  
7619 TGGGAATGAAAGATGAAATGCAACCTACCTCCCAAGAACCCAGAACTTATTAATAA 7678  
7693 agctctaagtgcccccgaatlaagaacctctgtcgtcgaatlaagtagtagaaggtca 7752  
7679 AACTTAATGCCCCCGAATTCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7738  
7753 caacttctatgttcccaagggcgtctatccctgagcctlaagtagaatacagaanaatgagtt 7812  
7739 CACTTCCATATTTCTCAGGCGCTGCTATCTGAGCTTAGTAAGATTAACAGAAATGAGTT 7798  
7813 gactaactcgttactcgtatctgttaaaactgactgagcagcaatagaaagaatgtattac 7872  
7799 GACTAATGCTTATCTGATGATTTCTGAAACAGTACAGGACCATTAAGAAATGATTAAC 7858  
7873 attgacagccctagtagactatctcaactgcaactgtcaactgtccagagagccacgc 7932  
|||||





QY 61 aagacgcttcgtgtgatttgggtgctgcctctccagaccggagagggat 120  
|||||  
Db 550 AAGACCCCTCTTGTGATTTGGGGTTCGGCTCTCGAGCCCGAGAGGGGAT 609  
QY 121 tttctttactgaccttcaatttgggttggccgggaaatccctgcacccctta 180  
|||||  
Db 610 TGTCTTTATGAGCCCTTCAATTTGGGCGTGGCCGGGAAATCTCGGACCAACCCCTTA 669  
QY 181 caaccggaacccgacttggaggtlaaaggaatccctcttgaacgtgtgtgtcggcc 240  
|||||  
Db 670 CACCCGGAACCGACTTGGAGGTAAGGATCCCCCTTTGGAAGCTGTGTGTGTGGCC 729  
QY 241 ggcgtctgttctgagtgctgttcttgcgtgagtcgcgcttcogtlttgcagctgct 300  
|||||  
Db 730 GGCCTCTATGTCTGAGATGCTGTGTTTCGGTAAATGCGCGCTTTCGGTTTSCACTGTCCT 789  
QY 301 ctacagacgttaagagacttggagagactgtatcaagacggtgtctagaagagatacaagctg 360  
|||||  
Db 790 CTCAGACCGTAAGAGACTGGAGGAGCTGTGATCAGACGCTGTTAGGAGGATCAACAGGCTG 849  
QY 361 ccaacccttgggagaccccgaggagtgaggagagccagagacgctgtgtctctact 420  
|||||  
Db 850 CCACCCCTGGGGAACGCCCGGGAAGTGGGAAGCCAGAGAACCCCTGTGTGTCTCTACT 909  
QY 421 gtgcgtacagagagacgagctgtgttltgaagcgaaagcttccctcccgcggtccg 480  
|||||  
Db 910 GTGCGTACAGAGGTCCGAGTCTGTGTTGTAAGCGAAAGCTTCCCTCCGCGGCGCTCGG 969  
QY 481 accttttgcctcttggagaagcgagcggtgcggtgtgtctggatctgtgtt 540  
|||||  
Db 970 ACTCTTTTGCCTGTGTGGAGAGCGGGAGCGGGGTGCGGTGTGTGTGATCTTGTGTGTT 1029  
QY 541 ctgtctcgtgtgtcttctgtctgtgtgcctcttctacagtttaatatagagacagag 600  
|||||  
Db 1030 CTGTTTGTGTGTCCTTGT 1089  
QY 601 tgaatacccccttagtttgaactctgcagacattgagctgaagtttagatccaggtctcata 660  
|||||  
Db 1090 TGACTACCCCTTGT 1149  
QY 661 attgtcagttcaggttaagaagagacgttgcagacttctgtcctcgtgaatgagcga 720  
|||||  
Db 1150 ATTGTGTAGTTACAGTTAAAGAGGACCTTGGAGACTTGTGTGTGTGTGTGTGTGTGTGT 1209  
QY 721 cattcgaatgttgaatgagcagagagagagagagagagagagagagagagagagagag 780  
|||||  
Db 1210 CATTCGATGTTGATGGCCATCAGAGGGGACCTTTAATCTGAAATTAATCTGTGTGTGT 1269  
QY 781 aggcacatcatlttccagactgagaccggtctcatccctgaltcagagagccctatalcccta 840  
|||||  
Db 1270 AGGCATCATTTTTCAGACTGAGACCGGGCTCATCTGTATCAGAGCCCTATATCTTAA 1329  
QY 841 cgtggaagaatttggcagaagaatccctccgcatgggttaacatggtctaaataaaca 900  
|||||  
Db 1330 CGGTGGCAAGTTTGGCAAGAAACCTCCGCAATGGGTAAACATGCTGCTAAATAACCA 1389  
QY 901 gaaagcagaatccccaatccctgtctcttggagagaaaaaacaacactgcgcccgaagaag 960  
|||||  
Db 1390 GAAAGCCAGGTCCCGCATCTGTGCTTGGAGAGAAAAACAACACTGCGCCGAAAAAG 1449  
QY 961 tgaagccctctctccgtatctaccccgagatcgaagagccgagacttgcgcggaacc 1020  
|||||  
Db 1450 TCGAGCCCTTC-TCTGTGTATCTACCCCGAGATGAGAGAGCCGCGACTTGGCCGGAACCC 1508  
QY 1021 caactgttcccccacccttatacagacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1079  
|||||  
Db 1509 CAACCTGTTCCTCCACCCCTTATCCAGACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1568  
QY 1080 ccttgaagctcagttgttgaag 1139  
|||||  
Db 1569 CCTGAGAGCTCCGGTGTGGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1628  
QY 1140 cggagagcgaacagagagatcgcatattaccgcttgcagcactatgtgcccctccatgtcca 1199

Db 1629 CCGAGACGGACAGACAGATCCGATATTACCGCTGCGGCACTATGAGCCCTCCATGCCA 1688  
QY 1200 gggggccaattgcagccctccagatattgcccctttcttcgaagatctcataattgg 1259  
|||||  
Db 1689 GGGGGCAATTGTGAGCCCTCCAGTATGTGGCCCTTTCTTCTGTGAGATCTATTAATGG 1748  
QY 1260 aaaactaacatccctctctccgagagatacccaacgctccagaggttggtagagctc 1319  
Db 1749 AAAACTAACATCCCTCTTCTTCGAGAGATCCCAACGCTACCGGGGTGGTAGAGTCC 1808  
QY 1320 ctatgtctctcaacagcctacttggagtagtgtcaacagctgtctgaagacatcttc 1379  
|||||  
Db 1809 CTTATGTTCTCACACAGGCTACTTGGATGATGTGTCAACACGCTCTGACACACTTTC 1868  
QY 1380 acaacgagagacgag 1439  
|||||  
Db 1869 ACAACCGAGAGACGAGAGAGATTTCTGTAGAGGCTAGAAAAAATGTTCTGTGGGCGCAG 1928  
QY 1440 gggcgacccacgaglttgcanaatagatctgacatggagatttcccttgcctgcgcccgt 1499  
|||||  
Db 1929 GGGCGACCCACCGCAGTTGCAAAATGAGATTGACATGGGATTTCCCTTACTGTGCCCGGT 1988  
QY 1500 tggagctacaacacggtctgaaggttagagagagagagagagagagagagagagagag 1559  
|||||  
Db 1989 TGGGACTACAAACAGGCTGAGAGGTAGGAGAGCTGAAATATCTATCGCAGGCTGTGGTG 2048  
QY 1560 ggcggtctcggggggggtcccaagagcgccactaatlttggctaaagtagagagtagtg 1619  
|||||  
Db 2049 GGGGGTCTCGGGGGCGCCCAAGACGCCCACTAATTTGGCTTAAAGAGAGGTGATG 2108  
QY 1620 caggagacggaacgaacctccctccgttattctcttgaagagctcatagagaccttgcagcg 1679  
Db 2109 CAGGGACCGAAGAACCTCCCTCGGTATTTCTTGAAGAGCTCATAGAACCTTCAGCGGG 2168  
QY 1680 ttcaccccttctgacttaacctagaagggccgaagacccctgaagcttgccttgccttca 1739  
|||||  
Db 2169 TTTACCCCTTTTGTGATCTTACCTCAAGAGGCCCGAAGAGCTGAGTGGCCCTGCTTCAT 2228  
QY 1740 gggcaggtgcgtctgatatatacagaaacttgcagaaacttgcagaaacttgcagaaact 1799  
|||||  
Db 2229 GGGGAGTGGCTCTGTGATATCAGAAAGAACTTCAAGACCTGGAAGGGTTCAGAGAGCT 2288  
QY 1800 gagttacgttgaatctagtag 1859  
Db 2289 GAGTTACGTGATCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2348  
QY 1860 gagaaggaacag 1919  
|||||  
Db 2349 GAGAGGAAACAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2408  
QY 1920 caagagagaatcttgaactaagatcttgcgcgagtggttgaagagagagagagagagag 1979  
|||||  
Db 2409 CAAGAGAGAAATTTGATCAAGATCTTGTGCGCAGTGTGTTGAAGGGAAGAGAGAGAG 2468  
QY 1980 aagagagagagatttagaagaaataggtcagagccctaaacagtcagagagagagagag 2039  
Db 2469 AGAGAGAGAGATTTAGAGAAATTAAGTCAAGGCTTAACACGTCAAGGAACTTGGGCAT 2528  
QY 2040 aggacccacttcagacagagccagltgtcgtatgttgaagagagagagagagagagagag 2099  
|||||  
Db 2529 AGGACCCCACTGACAGCAAGCAGGTGTGCTATTGTAAAGAAAAAGCACTGGGCAAGG 2588  
QY 2100 aacttcccccaag 2159  
Db 2589 AACTGCCCAAGAAAGAGAAACAAAGAGCCGAAAGGCTTCAAGGATTAAGATTAAGAT 2648  
QY 2160 tagggagagagaggttgcagacccctcccgagcccaagaggttaacttgaaggttgaaggg 2219  
|||||  
Db 2649 TAGGGAGAACGGGGTTCGAGACCCCTCCCAAGCCAGAGGTAACTTTGAAGGTGAAGGGG 2708  
QY 2220 caaccagttgagcttgcgttgaatacagagagagagagagagagagagagagagagagag 2279  
|||||



Dh 2709 CAACAGTGTGATCTGTTGATACGAGCGAGCATTCAGTCTGTACAAACATTA 2768  
Qy 2280 ggaactaaagaaaaaactcctgggtatgggtgcacagggcagcgagatcaca 2339  
Dh 2769 GGAAGAACTAAAGAAAAAATCTGGGTGATGGGTGCGACAGGCAACGGCATATCA 2828  
Qy 2340 tggactacccgaagaaacgcttgaacttggagtggaagtgaaacccctgttctgtc 2399  
Dh 2829 TGGACTACCCGAGAACCGTTGACTTGGAGTGGAGCGGGAAACCCAGCTGTTCTGTGC 2888  
Qy 2400 atcccttagtgccagatcccttctagtgtagagacttacttgacaaagatggagatcca 2459  
Dh 2889 ATCCCTAGTGTCCAGTACCCCTTCTAGTGAAGACTTACTGACCAAGATGGAGCTCAA 2948  
Qy 2460 attcctttgacaagaagacacagagtgctgtgataaacaacccatcactgtgtg 2519  
Dh 2949 ATTTCTTTTGAACAAGAGACCGAAGTGTCTGTGAATTAACAACCCATCAGTGTGTG 3008  
Qy 2520 accctccaatagaatgaatgaatataatctcccaagtaaaacccctgaatcaaat 2579  
Dh 3009 ACCCTCAATTAGATGATGAATATGACTATATCTCCCAAGTAAGCCCTGATCAAGAT 3068  
Qy 2580 atacaagctgtgttgagagcaggttcccaagcctggcagaacacagagatgggttg 2639  
Dh 3069 ATACAGTCTGTGTGAGCAAGTTCCTCAAGCCTGGGCAAAACCGAGGATGGGTGTG 3128  
Qy 2640 gcaagaagaagtlcccccacaggtatataactgaagacagtgctacacagatcacgc 2699  
Dh 3129 GCAAGAGCAAGTCCCGCACAGGTTATCACTGAAGGCCAGTGTATACCAAGATATCGTC 3188  
Qy 2700 agacagtaaccccttgatagaagagctcgagaagaatttggccgaatgttcaagaatta 2759  
Dh 3189 AGACAGTACCCCTTGAAGTGAAGGCTCGAAGAGAAATTTGGCCGATGTTCAAGATTA 3248  
Qy 2760 atccaacagggcatcctcagttcctgttccaaatcccttggatatactccctgcataccggt 2819  
Dh 3249 ATCCAACAAGGCAATCCTAGTTCTCTGTCCATCCCTTGAATACTCCCTGCTACCGGTT 3308  
Qy 2820 aggaagcctggagcaaatgatalcagacagtaacagacttgaagagtgtaataaag 2879  
Dh 3309 AGGAAGCCTGGAGCAATGATTTATCGACAGTACAGACTTGAAGAGGTCATTAAGG 3368  
Qy 2880 gtgagagacatacaacccaaggttccgaacctataactcttggagcgccctccgct 2939  
Dh 3369 GTGAGAGACATACACCAAGGTCGCCAACCCTTATTAACCTTTGACCGCCCTCCGCCCT 3428  
Qy 2940 gaaaggaactgtacacaglatitgaacttaaaagatccttcttcccgagatlaaac 2999  
Dh 3429 GAAGGAACTGCTACACAGTATGACTTAAAGATGCCCTTCTTCTGCTGAGATTACAC 3488  
Qy 3000 ccaactagccaacacttttgccttcgaatggaagatccaggtacgggaagaccg 3059  
Dh 3489 CCCACTGAGCCAAACCTTTTGGCTTGAATGAGAGATCAGATGAGGGAACACCGGG 3548  
Qy 3060 cagctaacctggagccgagctggcccaaggtttaaagaactcccgaccatcttggacaa 3119  
Dh 3549 CAGCTCACCTGGAGCCCACTGCCCAAGGCTTAAGAACTCCCGAATTTCTTTGAGAA 3608  
Qy 3120 gccctacagagagcctgccaacttcaagatccaacacccctcaggtgacccctccag 3179  
Dh 3609 GCCCTACACAGGAGGACCTGGCCAACTTCAGATCCAAACCCCTCAGTGAAGCCCTCTCCAG 3668  
Qy 3180 taagtgatgaactgtcttgcgggagacccaacaggaactgtctagaagtlacgaag 3239  
Dh 3669 TACGTGATGACCTGCTTCTGGGGGAGCCACCAACAGSAGTCTAGAAAGGTACGAAG 3728  
Qy 3240 gcaactactgtcgtgaattgtctgacctaggttacaagagcctctgctaaagagccagatc 3299  
Dh 3729 GCACTACTGCTGGAATGTCTGACCTAGCTACAGACCTCTCTAAAGAGCCAGATT 3788  
Qy 3300 tgcagagagaggaatacacttgggttacaagtttgcgggagggagcagatgctgagc 3359  
Dh 3789 TGCAGAGAGAGTAACATACTTGGGGTACAGTTTGGGGGGGGGAGCATGTGCTGACG 3848  
Qy 3360 gaagcaggaagaaactgtatccagataccggccccaacacagccaagtgaga 3419  
Dh 3849 GAGGCAAGGAGAGAAACTTACTGATCCAGATACCGGCCCCACACCAACCAAGTATGAGA 3908  
Qy 3420 gaatttgggagacagctgagatlttgacagcttgatccggggttggagacttaaga 3479  
Dh 3909 GATTTTGGGGACACTGGATTTTGCAGACTGTGGATCCGGGGTTTGGCAGCTTAGCA 3968  
Qy 3480 gcccaccttaccggcttaacaaagaagaaggatctcctggctccttagacag 3539  
Dh 3969 GCCCACTTACCCGCTAACCAAGAAAGGGAATTCCTGGGCTCTGAGACAG 4028  
Qy 3540 aaggaatttgatgatacaaaaagccctgcgcgaagcagcagctgtgcccctctgac 3599  
Dh 4029 AAGGCAATTGATGCTATCAAAAAGCCCTGCTGAGCGCACCTGCTGCTCCCTGAC 4088  
Qy 3600 gtaactaaaccttaccctttagtgaagcgtlaaaggagtagccagagagttta 3659  
Dh 4089 GTACTAAACCTTTTACCTTTATGTGAGCGCTTAAGGAGTGAAGCCGAGAGATTTTTA 4148  
Qy 3660 acccaaaccttaggaacatggagagacctgttgcctactgttcaagaagcttactct 3719  
Dh 4149 ACCCAACCTTAGAGACATGAGAGACCTGTTGCTTACCTGTCAAGAGCTGATCT 4208  
Qy 3720 gtagccagtggttggccgtatgltcgaaggtatcagcagctgttggcactgttcaag 3779  
Dh 4209 GTAGCCAGTGTGGCCGTATGCCGAAGCTTATCGACGCTGTGCCATGTGTCAG 4268  
Qy 3780 gacgttgaacaattgaacttggagacaataaactgtlaaataagccccatgcatltgag 3839  
Dh 4269 GACGCTGAACAAATTTGATTTGGGACAGATATATATGATGATGATGATGATGATGATG 4328  
Qy 3840 aacatgttgcgcagcccccagacagatgtagtgaacagcccgatataccacataca 3899  
Dh 4329 AACATGTTTGGGAGGCCCCAGACCGATGATGATGATGATGATGATGATGATGATGATG 4388  
Qy 3900 agcgtgtcttcaacagaagaggttcaacttgcctcacaacagccgctcacaacctgac 3959  
Dh 4389 AGCCTCTTCTCAGAGAGAGGCTCACTTTCGCTCCACAGCGGCTCTCAACCTGCACT 4448  
Qy 3960 ctttgccttgaagagacttgaacacagtgatctatgttgcatacaactatgttgaag 4019  
Dh 4449 CTTCCTGCTGAAGAGACTGATGATGATGATGATGATGATGATGATGATGATGATG 4508  
Qy 4020 ggaacttgggttcgcaagagaccttacaacacacagctgaactggagagtgtaacctg 4079  
Dh 4509 GAGACTGGGGTCCGAGAGGACCTTACAGACATACCGCTGAGTGAAGTGTGTAACCTG 4568  
Qy 4080 ttaacttgaagagagacttgaatgttgaaggaaggaaggaaggaaggaaggaaggaag 4139  
Dh 4569 TTCACAGAGAGAGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATG 4628  
Qy 4140 gacggagacccgagacatctgggacagccgctggcgggaaggaacttgaagcgaagag 4199  
Dh 4629 GAGGGAGCCGACAGATCTGGGGCAGACGCTGCGGGAAGGAATTCAGCGCAAGAGCT 4688  
Qy 4200 gagctatggccctcaagcaagcttggcgttgcgcgaaggaaggaatccaataaacttat 4259  
Dh 4689 GAGCTATGGCCCTCAGCGAAGCTTTGGGGCTGGGCGAAGGGAATTCATTAACCTTAT 4748  
Qy 4260 acggaagcaggtatgtccttggagactgaacgltacacggggccaactataaacaag 4319  
Dh 4749 ACGGACAGAGGTATGCTTTTGCAGCTCACACGTACACGCGGCGCATCTATTAACAAAG 4808  
Qy 4320 ggggttgaactcagcaggggaggaataaagaacaaagggaattcctaagctata 4379  
Dh 4809 GGGTGTCTTACTCAGCAGGAGGGAATTAAGAAAGGAAGGAATTCCTAAGCTTATTA 4868  
Qy 4380 gaagccttacttgcacaaagagctatatacactgtcctggagacacagaagcc 4439  
Dh 4869 GAAAGCTTATGATTTGCAAAAAGGCTTACTATATATACCTGTCTGGACATCAGAAAGCC 4928

QY 4440 aaagatcctcatctagagggaaccagatgctgaccgggtgtgccaagcagcccg 4499  
|||||  
Db 4929 AAGATCTCATATCTAGAGGGAACCAAGTGGCTGACCGGGTTGGCCAAAGCAGGACGCCAG 4988  
4500 gctttaacctctgctcctaataagaacgcccagaagccccagaaccagagacgtac 4559  
|||||  
Db 4989 GCTGTTAACCCTTGGCTATATATAGAACGCCCAAGCCCCCAAGCCCCAGACACAGTAC 5048  
4560 accctagaagaactgcaagaagataaaagatagaccagttctctgtagactccgagggg 4619  
|||||  
Db 5049 ACCCTAGAAGACTGGCAGAGAGATATAAAGATAGACCAGTTCCTCAGACTCCGGAGGGG 5108  
4620 acctgtatactcctacataggggaagaatctgcccacaaagaaggttagaatactc 4679  
|||||  
Db 5109 ACCGTGTATACCTCATATATGGAAGAAATCCTGCCCAAGAAAGGTTAGATATGTC 5168  
4680 caacgatacatcgtctaaccccactagggaactaaacacctgagagttgtgtcagaaca 4739  
|||||  
Db 5169 CAACAGATATCATGCTCTCACCCACCTAGAGAACTAAACACCTGCAGAGTTGGTCAGAAACA 5228  
4740 tcccctatcatgtctctagaggtctaacaggagtgtgctgactcggtgtcacaactgtgt 4799  
|||||  
Db 5229 TCCCTTATCATGTTCTGTAGGGCTACACGAGAGTGGCTGACTCGGTGTCAAACATTGTGTC 5288  
4800 cccgtccagctgtgttaatgtctaactcctccagaatacctccaggaagaagactaaaggga 4859  
|||||  
Db 5289 CCTGCACAGCTGGTTATGCTATCTTCAGAAATACCTCCAGAAAGAGACTAAGCGGA 5348  
4860 agccaccocggcgctcaactcggggaagtgtgacttcaactgaggttaagccggctaaatcggga 4919  
|||||  
Db 5349 AGCCACCCAGGCGCTCAGTGGGAAGTGGACTTCAGTACGAGTAAAGCGGCTAAATACGGA 5408  
4920 aacaatatctatgtgtttttragaacaccttccaggaatggtagaggcttatccctact 4979  
|||||  
Db 5409 ACAAATATCTATTTGGTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCTTACT 5468  
4980 aagaagaagaactcaacccgtgtgtcgttaagaanaatacttgagggaatlttcccaagatt 5039  
|||||  
Db 5469 AAGAAGAGACTTCAACCGGCGTGAATAAATAATCTGAGAGAAATTTTTCGAGATTT 5528  
5040 ggaataccctaaagtaataggttcagaacatggttccagcttctgttggcaggtaagtcag 5099  
|||||  
Db 5529 GGAATACCTAAAGGTAATCGGGTCAGCAATGGTCCAGCTTTTGTGCCAGGTAAAGTCAG 5588  
5100 ggaactggccaagatatgggagatgtatgtgaacatgcatgtgtcataagagaccccaaac 5159  
|||||  
Db 5589 GGACTGGCCAAGATATTGGGGGATTTGATTTGAAACTCTCATTTGTACAGACCCCAAGC 5648  
5160 tcagaacaggtagagaagaatgaatagaaccatlaaagaagacccttaaccaattgacaca 5219  
|||||  
Db 5649 TCAGGACAGGTAGAGAGATGAATAGAACCATTAAGAGAACCCCTTACCAATTTGACACACA 5708  
5220 ggaactggcaatlaatgtatgtatggtctcctgccccttgtgtcttttaggttgaagac 5279  
|||||  
Db 5709 GAACACTGGCATTAATATTGATGATGATGCTCTGCCCTTTGTGCTTTTGGGTTAGAAC 5768  
5280 accctgggaacagtttgggtcgaaccccttaagaattgcttaagggggagacccccccgttg 5339  
|||||  
Db 5769 ACCCTGGACAGTTTGGGCTGACCCCTTATGAATTCCTTACGGGGGACCCCCCGGTTGG 5828  
5340 gcaagaattgaccttgacacatagtgctgaatgtgtcgtcttccagagcttgttcttag 5399  
|||||  
Db 5829 GTAGAATATGCTCTGTACATAGTGTGATGTGCTGTTCCOAGCTTTGTCTCTAG 5888  
5400 ctcaagggcgtcagtgaggttgaagcagcagcgttgaagcagctcccggaaggcctaacta 5459  
|||||  
Db 5889 CTCAGAGCGCTCAGTGGGAGGAGCAACGCGTGAACACCTCCGGAGGCGCTACTACA 5948  
5460 ggaagacttgaagcttcaacatgcttccaaagtggagattcagctcatgttga 5516  
|||||  
Db 5949 GGAAGAGAGACTTTCAGAGTCCACATCCGCTTCCAAAGTGGAGATTCACTATATGTTAGA 6008  
5517 cgcaccgtgtcaggaacacctcgagactcgttgaaggaacctatctgtactttagacc 5576

Db 6009 CCCACCATGCAAGGAACCTCGAGACTGGGGAAGGGCCCTTATCTGTACTTTTGACC 6068  
5577 acaccaagcgtgttgaagtcgaagaaatccccccttgatccatgacatcccaagaa 5636  
|||||  
Db 6069 ACACCAAGCGCTGTAAAGTCAAGGAATCTTCACCTGGATTCATGCAATCCACGTTTAA 6128  
5637 ccgagccaaactccccaactcgggtgtgagagccgaagaagacttgaatcccttaagct 5696  
|||||  
Db 6129 CTGGGCGCACCTCCGACTCGGGGTGAGAGCCGAAGAAAGACTGAGAATCCCTTAAGCTT 6188  
5697 cgcctccatcgctgtgttcttacttacttaacaataactcccaagcagtagtaaacgct 5756  
|||||  
Db 6189 CGCCTTCATCGCCTGGTCTTACTCTAACAATACTCCAGGCGCACTAGTAAACGCTCT 6248  
5757 tatagaagcttcgaaccccccaataagaccttacccttaaccggtgattatggccctga 5816  
|||||  
Db 6249 TATAGACAGCTGGAACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTTGACCTGA 6308  
5817 tacagggtcactgtaaaatagacactcggaggtgtgtcctcttagagagcactgtgtcga 5876  
|||||  
Db 6309 TACGGGTGCATGTAATATAGCACTCGAGGTGTCTCTAGAGGCACTGGTGGCTTGA 6368  
5877 actgacttctgctcgcattgattaaacccgctgtttaaagaacacctcccaactagt 5936  
|||||  
Db 6369 ACTGCATTTGCGCTCCGATTGATTAAACCCGCTTTAAAGACACACTCCCAACTAGT 6428  
5937 ccgtatgtataggttcttacttctgcccaggaacagagaagaagaatctgttgggtctc 5996  
|||||  
Db 6429 CCGTATATATGGTCTATTTGCTGCCAGGACACAGAAAGAAATCTGTGGGGTTC 6488  
5997 tgggaatcctctctgtaagagatgtagctgtgtaacctccaaagatgtagagctgaatg 6056  
|||||  
Db 6489 TGGGAATCTTCTGTAGAGATGAGAGTCCGTCACTCCACCATGAGAGCTGGAATG 6548  
6057 gccgaatcctctccaggaaccgggtaaatctccttltgtcaatcccgccggcgaagta 6116  
|||||  
Db 6549 GCCGATCTCTCCAGGACCGGGTAAATTCCTTTGTCAATTCGCCGCCGCGCAAGTA 6608  
6117 caaatgtatgaacctataataagaatagctgtcccatcagacttaagtatctaaa 6176  
|||||  
Db 6609 CAAAGTGAATGAACATATATTAAGATTAAGGTGCTCCCACTCAAGACTTAAGTATCTTAA 6668  
6177 gataagttcaactg--aaaggaacaaagaaatalttaaaagtggataagtgtag 6234  
|||||  
Db 6669 GATAGTTTCACTGAAAAAGGAAAAACAGAAATTTTAAAGTGGATTAATGGTATGAG 6728  
6235 ctggggaatgttttat-tatatgcyggggagcaggtgccacttaaccatctgct 6293  
|||||  
Db 6729 CTGGGGAATAGTTTATTAATATGCGGGGAGAGAGGTCACCTTAAACCATTCGCT 6788  
6294 tagatlaagacgggggaacgaacccccgtgtgcaatgggagccgaataagtaactgctga 6353  
|||||  
Db 6789 TAGATTAAGAGAGGGGAGAGAAACCCCTGTGGAGTGGAGCCCATTAAGTAAGTGGCTGA 6848  
6354 acaagggcccccggcctggagcaccgcatcaacttgcggtggcccaatcaactcgt 6413  
|||||  
Db 6849 ACAGGGGCCCGCGCGCTGGAGCCAGCATTAACCTTGGCGGTCGCCAATTAACCTGCT 6908  
6414 gcygcttgacataaacacacgcgcttagcaacagtaaccaacttgaatttccatacaaac 6473  
|||||  
Db 6909 GCGGCTGACATTAACACAGCGGCTAGACANAGTACCACTGGATGATTCCTACCAACAC 6968  
6474 gcttagaaactcccaaggtgttctgtttaaagaagaagaagactcttagtccatcca 6533  
|||||  
Db 6969 GCTTAGAAATCCCGAGGTGTTCTGTTTAAACAGGAGAGAGACTCTTCACTTCATCCA 7028  
6534 ggaagcttccaaagcactcaactcaacgaccggatgacacttcttctgttggcttg 6593  
|||||  
Db 7029 GGAAGCTTCCAGGCACTCAACTCCAGGACCTGATGCCACTTCTTGTGGCTTGG 7088  
6594 tctatcctcaggtcctcttattatgagggatgagctaaagaaanaaatcaatgtgac 6653  
|||||

Db 7089 TCTATCTCAGGGCCCTCTTTATGAGGGATGGCTAAGAGAGAAAATTCATGTGAC 7148  
 Qy 6654 caaagaagcatagaatacaatgacatgggggtccagaataagcttaacccactgaagt 6713  
 Db 7149 CAAGAGCATAGAAATCAATGATACATGGGGTCCGAAATAGAGTTACCTGACTGAGT 7208  
 Qy 6714 ttccgggaagggagacatgacatagaataaagctcccccatcccaacaacacttgctatg 6773  
 Db 7209 TTCCGGGAAGGGAGCATGATAGAGAAAAGCTCCCCCATCCCAACACCTTTTCCAAATAG 7268  
 Qy 6774 tactggtttatagcagcgcctcagaataatcagatattagttacctggttaacaagtg 6833  
 Db 7269 TACTGTGGTTATGAGCAGGGCTCAGAAAATCACTATTATAGTACCTGTTATACAGTG 7328  
 Qy 6834 gtgggcatgcaataactggttlaacccctgtgttccacactcagttccacaatccaa 6893  
 Db 7329 GTGGGCATGCAATACTGGGTTAACCCCTGTGTTCCACCTCACTTTCAACCAATCCAA 7388  
 Qy 6894 agatttctgtcatggtcacaatcgtcccccagtgactacatcctcaggaagtggt 6953  
 Db 7389 AGATTTCTGTGTCATGGTCCAAATGTCCTCCGAGTGTACTACCATCTCGAGAGTGT 7448  
 Qy 6954 ccttgatgaatagactacatcgtglaaacccgaacaaagaagacccglatcccttaccct 7013  
 Db 7449 CCTTGATGAATATGACTATCGGTATTAACCCGCAAAAAGAGAACCGTATCCCTTACCT 7508  
 Qy 7014 agctgtaactcgtgattaggaagccggtgtgctgtaggaacaggaacagctgcctgat 7073  
 Db 7509 AGCTGTAAATGCTCGATTAGGAGCGCGGTGGGTAGGAACAGGGACAGTCCCTGAT 7568  
 Qy 7074 cacaggacacacagcagcagtagaagaagactggtgactaactcagccatcagagaaga 7133  
 Db 7569 CACAGAGACACAGACGCTAGAGAAAAGACTTGTGAGCTACATGCGGCATGACAGAGA 7628  
 Qy 7134 tctccgagccttaagaagagctgtttagcaaccltagaagagctccctgactcttctgta 7193  
 Db 7629 TCTCCGAGACCTTAAAGAGGAGTCTGTAGCAACCTAGAAAGAGTCCCTGATCTTGTCTGA 7688  
 Qy 7194 agtgggttctacagaacccggagggagtagatcgtcgttcttaagaagaagtggtgatg 7253  
 Db 7689 AGTGGTCTTACAGAACCGGAGGGATTAAGATCTCTTTCTAAGAAAGGTGGGTATG 7748  
 Qy 7254 tgcagccttaagaagaagtgctgtctctatgtagatcactcaggaagccatcagaagctc 7313  
 Db 7749 TGCAGCCTTAAAGAAATATTCCTTATGTATGATCAGTCAAGAGCCATCAGAGACTC 7808  
 Qy 7314 catgaacaagcttagaanaaagcttagaagagcgtcgaagggaggaagagagcctgaaccagg 7373  
 Db 7809 CATGAGCAAGCTTAGAGAAAGTTAGAGAGCGCTCGAAGGGAAAGAGAGCTGACAGGG 7868  
 Qy 7374 gtggtttgaagaagatggttcaaacaggtctccttgtagatgacacacccctgctcgtgac 7433  
 Db 7869 GTGCTTTGAAGAGATGTTCAACAGGTCTCTTGATGACCAACCTGCTTCTGCTCGAC 7928  
 Qy 7434 gggggcccttagtagtctcgtcctgttactacagttgggccttgcttaataatagtt 7493  
 Db 7929 GGGGCCCTTACTACTCTGCTCCTGTTACTTACAGTTGGGCTTGGCTTATTAATAGTT 7988  
 Qy 7494 tgtgctcttgtagaagacagatgagtgacatgcacatcatgtagttagttagcaacagta 7553  
 Db 7989 TGTGCTTGTGTAGAGAAAGAGTAGAGTGCAGTCCAGATCATGTACTTAGGCAACAGTA 8048  
 Qy 7554 ccaagggccttctgagccaagaggaagaactgacacttagccttccagttcttaagattaga 7613  
 Db 8049 CCAAGGGCTTCTGAGCCAAAGAGAACTGACCTTACCTCCAGTTCTTAAGATTAGAA 8108  
 Qy 7614 ctattacaagaagaagaagtgaggaaatgaaagataagaaatgaaacttaacccctccaga 7673  
 Db 8109 CTATTAAACAGACAGAGATGGGAAATGAAGATGAAGAACTTAACCTCCAGCA 8168  
 Qy 7674 acccagaagttataaagaagcttaaatgcccccgatatacagaacccctgctgctgca 7733  
 Db 8169 ACCCAGAGATTATAAAGCTCTAATATGCCCGCAATTCAGACCCCTGCTGCTGCCA 8228

Qy 7734 gtaaatagtagaaggtcacacttccatattgttccaggccgtgctatccctgagtaagta 7793  
 Db 8229 GTAATATAGTAGAAGGTCACACTTCTGATTTGTTCCAAAGCCGTGCTATCTGGCTAAGTA 8288  
 Qy 7794 agatacagaagaatagtgtagtaactcgtctatcgtgattcgttaaaactgactggcacc 7853  
 Db 8289 AGATTAACAGGAATATGATGACTTAATCGCTTATCTGGAATTCGTAAACCTGACTGGCAC 8348  
 Qy 7854 atagaagaattgattacacatgacacccctagtgacatcctacatcgaactcgtcac 7913  
 Db 8349 ATAGAAATATGATTTACATTGACAGCCCTAGTACCTATGTGAACCTCAATCTGTGTGAC 8408  
 Qy 7914 tctgcccaggaagcccaacgaagatgagacccctccgagactatttaaatgattgttccac 7973  
 Db 8409 TCTGCCAGTAGAGCCACGACGATGCGGACCTCCGAGCTATTTTAAATGATTTGGTCCAC 8468  
 Qy 7974 gtagggcgggctcgtgatttttaaatattgtgttcacagggagccgggctcgtatatt 8033  
 Db 8469 GGAGCCGGGCTCTCATATTTTAAATATGATGTTCCATGAGAGCGGGCTCTCATATTT 8528  
 Qy 8034 ttaaatgattggttgttgaacgacagcttgttgaacccatlaaagaactgtccga 8093  
 Db 8529 TTAATATGATTTGTTATGACAGCGACTTTGTTGTGAACCCCAATAAAGCTGTGCCGA 8588  
 Qy 8094 ttccgactcggggccgacgctcctacccctggtgtgtgtagacagctgtggcccaagc 8153  
 Db 8589 TTCCGACATCGGGCGCGCATCTTACCTGCGGTGGTGTACGACTGTGGGCCCCAGCG 8648  
 Qy 8154 cgttggaatlaaaatcctccttgcgttctgcatcgaaga 8191  
 Db 8649 CGCTTGGAATTAATATCTTGTCTGTGTTGATCATAAGA 8686  
 RESULT 4  
 PEN133818  
 LOCUS  
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class B, clone 43).  
 ACCESSION AJ133818  
 VERSION AJ133818.1 GI:6688949  
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM Porcine endogenous retrovirus.  
 REFERENCE 1 (bases 1 to 8750)  
 Czauderna, F., Fischer, N., Bolter, K., Kurth, R. and Tonjes, R.R. Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)  
 2 (bases 1 to 8750)  
 Tonjes, R.R.  
 Direct Submission  
 Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY  
 FEATURES  
 source  
 1. 8750  
 /organism="porcine endogenous retrovirus"  
 /proviral  
 /strain="type C"  
 /db\_xref="taxon:61673"  
 /clone="43"  
 1. 629  
 LTR  
 misc.feature  
 467  
 /note="cap site"  
 primer\_bind  
 632..649  
 /note="g1y4"  
 1064..2635  
 gene  
 /gene="gag"

mat\_peptide 1064. .2635  
/gene="gag"  
/function="retroviral capsid"  
/product="Gag protein"  
Join(1064. .2635,2639. .6223,6099. .8072)  
/note="viral polypeptide"  
/protein\_id="CAB65341.1"  
/db\_xref="GI:568950"  
/translation="MGQYTPPLSLDHWTEVRSAHNLISQVKKGRWTFCAEMF  
TFDVGWSEGTNSLITLAVKALIFQIGSSHPDDEPIILTMQDAEDPEPWKFWLN  
KPRKDPRLILAGEKNKSAEKEPSPRIPELPPWPEPVPYQVQVAGV  
PSAPGAPVVEGPAAGTRRGATERTDEILILRLYGPMPGQLOPYMPS  
ADLYNKTNNPESDPCRLTGIVSLMSHOPMDCCOLQTLFTTEERLIL  
RKNVPGADRPOLONEIDMGFPLRPGMDYATGARESIKYLRAIYAGRSAR  
TNLAKYREVMQGPNEPSPYFLERLMAERFPTPEPTSAQASALAFIGSALDIR  
KTLQRLGQLEALNDLVREKVRRTRETEKEDQREKEEREERDRKQENLT  
KILAAVEGKSSREERDRKTRSGRQSGNLNPTLDKDCACAKGHWANPK  
KNGKGRVALLEEDKGRGSDPLEPRVALKEGQPELVDTGAHSVLQPLGK  
KEKSWMGATGOROYPTTRTVDLAVGRHSFLVPECEVPLLRDLTKMAOI  
SFGQRPVYVNNKPTTYLTLOLDEVRYSPOVKPDODIOSMLRPOQAAETGMC  
LAKQRPVYIOLKASATPYVROQYPLSRARCTMPHORLIQOQITLVDSWPNPTPL  
LPRKPGINDYRPVODLREVNKRVDIHTVBNPNLSALPERNWTYVLDKDAFL  
CLRHPISQPLAFEMRDGRTGQLWTRLPQFKNSPTLFDALHRLDLNRIQH  
POVTLQYVDDLLAGATKODLEGTKALLLESLGYSASAKAQLCRRETYLGS  
LRGQRMLEARKKTYVOIPAPTTAKOVEEFLGTGFCRLMIPGATLAAPLYLTK  
KEESAPRKHOTPAIKKALLSAPALAPDYTKPTTYLDERKGVAVLTOTLGP  
RRVATLSKLDLPVAGSPVCLKATANAIVLKDKDKLTGNTIYAPHALENVRO  
PPDRMNTNARKHYSLLTERVTAPRALNPATLLPEIDEVTHDHOULLIETG  
VRDLIDILPTEBVLWFTDSSYVEGRMAGAVDGTHTIMASSLEPESAKAE  
LMAITOLALAEKGSINITYDSYAFATNHAIGAYIKORGLTSGRETKNEELISL  
LEALHPRLALIHCGHOKAKDLISRCNOMADRAKAOAVNLPIITETKAPER  
ROYTLEDMOETIKKIDOFSETPGTCYSKELHPKEGLEVOQIHRHLTGHTILO  
OLYRSYVHLIPGADVSVKCYPCOLVYMNAPRPVPGKRLRSGREKLEHNDTE  
VPRAKGKYLIVFDTSQWYATPTKEISTVYAKKLEIETPRRGTIPKVISDNG  
PAVAVOSQGLAKILGIDMKLHCAYRPOSSQGVEMNKTKETLTKETETGNIWA  
LLEPVLFRANTPGOGLPYELLYGGPPLVEISVHADVLISQPLSRKALAEW  
KORAMKOLREAYSGEDLOVPHRFQVGDVYRRRAGLLETMRKGPYLVLTPTAV  
KEGISTWILHSHVKLAPPDSGMAEKENPLKRLHLVLYSNNSGOMHPTLW  
RHLPRGGEPRKLRIDIPASIAMETLITPOASSKRLIDSSNRPCLSLWLTIDP  
DTCVYVNSTRGVAPRGTWPELHFCRLINPAVKSTPRLVSYSTPCPGTEKRYC  
GSESEFCRMSCVTSNDGDKWPIISLDORVFSFVNSPGYKWKYIKDSCSPSD  
LDYLKISTEPKROENIOKIMNGSMWIVFYVGGAGSTLIRLIEGEPVAVG  
PDKVLAEGPPLALEPHNLVPQLISLIPDIYTOPSNGTGLIPINTRNPSGVAKI  
GOLFSLIGAFQAINSTDPAISQMLTSSGPTVEYEMAKGPNYTKERNCOTW  
GSRNKLITLEVSGKGCICAKAPSHOHICYSVVYEQASENOYLXVGRNRMATCGL  
TPCVSSSVFNOSKDFCVMOIYPRVYIYHEEVLDEIRINRPREVSLTLVAMLG  
LGRVAVGVTGIALIITGPOLEKGLGEIHAAMTEDLRLAESVNLIESLISLSEVL  
ONRGLIDLFLREGGICAALEKCECCFYVDHSGAIDRMSKLEERLERREREADOGW  
FKGMFRSPMNTLLSALGPLVILLLLLTGVPCLINRFAVRERVSAVOIMVLROO  
YQGLISQGETLL"  
2639. .6223  
/gene="pol"  
/function="retroviral protease and polymerase"  
/product="Pol protein"  
2639. .6223  
/gene="pol"  
/gene="env"  
/gene="env"  
/function="retroviral envelope"  
/product="Env protein"  
6099. .8069  
/gene="env"  
/gene="env"  
LTR 8122. .8750  
BASE COUNT 2288 a 2214 c 2286 g 1962 t  
ORIGIN

Query Match 97.6%; Score 8011.6; DB 14; Length 8750;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 8124; Conservative 0; Mismatches 54; Indels 20; Gaps 6;

QY 1 gtcgtgtaacagctgcyggccccccagcgcctgtggaataaaatcctctgtctgtgtaac 60  
|||||  
Db 489 GGGGAGTACAGACTGGGGCCCGCCCTTGATTAATAAATCTCTGTTGATTCATC 548  
QY 61 aagaccgctctcgtgagtgattgagtgctctctcgaagccggaagggagat 120  
|||||  
Db 549 AAGACCGCTTCGTGAGGATTTGGGGGTGCTGCTTCCAGCCCGGAGAGGGGAT 608  
QY 121 tctcttcttaacgaccttcatcttggtcgttgcgcgggaaatccttggaacacctta 180  
|||||  
Db 609 TGTCTCTTTACTGGCCTTTTCATTTGGTGCTTGCCGGGAAACCTGCGACACCCCTTA 668  
QY 181 caaccgagaaacagactggaagtaaaaggaatccctcttgaaacgtgtgtgtgcgagc 240  
|||||  
Db 669 CACCCGAGAACCGACTTGGAGGTAAAGGATCCCTTGAAACGCTGTGTGTGCGGCC 728  
QY 241 ggcgtctctgtctgagtgctgttctcggtgagtcgcgcttgcggttgcagctgtctc 300  
|||||  
Db 729 GCGCTCTCTGTTCTGAGTCTC-----TTCCGCGCTTTCGTTTGACAGCTGCTCT 776  
QY 301 ctcagaacgtaagagacttgagactgtatcagaagacgtgtctaggagagatcaagagctg 360  
|||||  
Db 777 CTCAGACCTTAAGACTGGAAGACTGTGATCGACAGCTGTAGAGAGATCACAGCTG 836  
QY 361 caaccctggagagacgcccccgagaggtgggagagcaaggaagcctgtgtgtctact 420  
|||||  
Db 837 CCACCTGGGGGAGCCCGCGGAGGTGGGGAAGCAGCAAGGAGCCTGTGTGTCTCTACT 896  
QY 421 gtcggtcagaagagccgaagctgtctgtgtgtaagcgaagcttccctcgcgcgtctcg 480  
|||||  
Db 897 GTCCGTGACAGAGACCGAGTCTGTTGTGAACCGAAACCTTCCCTCGGCGCTGCG 956  
QY 481 actcttcttcgctctgtggaagagcgcgcgcgcgtgtctgtgactgtgtgtt 540  
|||||  
Db 957 ACTCTTTTCCCTGCTGT 1016  
QY 541 cgtctcgtgtgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtct 600  
|||||  
Db 1017 CGTTTTCGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT 1076  
QY 601 tgaactaccccccttgattgactctcgcgaacattggaactggaagttgataccaggtcata 660  
|||||  
Db 1077 TCACGACCCCTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136  
QY 661 attctcgaatgaagttgaagaggaacattggaacatttgcgtcctgaatggccaa 720  
|||||  
Db 1137 ATTGTCACTTGAAGTTAAGAGGACCTTGGCAACATTTCTGTCTCTGTAATGGCCAA 1196  
QY 721 catcgaatgtgaaatgagcatcagaagaggaacattgaatctgaaatatacctgtgtgta 780  
|||||  
Db 1197 CATTCGATGTGATGGCCATCAGAGGGGACCTTAATTCTGAATTAATCTGGCTGTTA 1256  
QY 781 aggcatacaatttcaagactggaacccgcgtctcatctcgatcagaagccataatcctta 840  
|||||  
Db 1257 AGGCAATCATTTTTCAGATGTGACCCAGCTGCATCTCATATGAGAGCCCATATATCTTA 1316  
QY 841 cgtgcaagatttggcaagaagatcctcgcgaatgggttaaaacacagtcggtcaataaaca 900  
|||||  
Db 1317 CGTGGCAACTTGGCAGAGATCTCTCCGCCATGGGTTAAACATGAGGCTTAATAAACCA 1376  
QY 901 gaaagcgaagtcgccgaatcctgctctgtgagagaaaaaacaacactcgcgcgaaaaag 960  
|||||  
Db 1377 GAAAGCAGAGTCCCGAATCTGCTGTGTGAGAAAAAACAACACATCGGCGCAAAAAG 1436  
QY 961 tgaagccctctcctcgtatcaccacagatcagagagccgcgaacttgccggaaccc 1020  
|||||  
Db 1437 TCGAGCCCTC-TCCTGTATCTTACCCCGAGATCGAGAGCCCGCATTTGGCGGAACCC 1495  
QY 1021 caactgttcccccccccttatcagacagaggtgtgtgtgagagga-cctctccccct 1079  
|||||  
Db 1496 CAACCTGTTCCGCCACCCCTTATCCAGACAGGAGGTGCTGTAGAGGACCCCTCTGCCCT 1555  
QY 1080 cctggaagctcgcgtgtgagaggaacctctgcgcggaactcgaagccggaagagcgcaacc 1139

|||||  
Db 1556 CTTGGAGCTCCGGTGGAGGAGGACCTGCTGCCGGGACTCGGAGCGGAGAGCGCCACAC 1615  
Qy 1140 ccgagacgagacagacagatcgcgaatatcaacgcctgcgcacatacgcctccatgcga 1199  
Db 1616 CCGAGCGGAGACAGAGAGATGCGATATTACCGCTCGGAGCTATGCGCTCCCATGCGG 1675  
Qy 1200 gggggccaatcgcagccctccagatattggcccttctctcgcagatcctataatgg 1259  
Db 1676 GGGGGCAATTCAGCCCTCCAGATTGGCCCTTTCTTCGCGAGATCTCATTAATTGG 1735  
Qy 1260 aaactaacccatcccccttccgagagatccccacgcctcagaggttggtagagcc 1319  
Db 1736 AAAACTAACCTACCCCTTTCTTCGAGAGATCCCAAGCCTCAGCGGTTGGTAGAGTCC 1795  
Qy 1320 ctatgcttctcaccagcctacttggagatgtgtcaacagctgcgcacacactcttc 1379  
Db 1796 CTTATGTTCTCTCACAGGCTACTTGGAGATGATTTGTCACACAGCTCTCAGACACTCTTC 1855  
Qy 1380 acaaccgagagcagagagaaatctctgttagagcgtagaaaaatgcttccctgggagc 1439  
Db 1856 ACAACCGAGAGAGAGAGAAATTCGTAGAGGCTAGAAAAAATGTTCTGGGCGCCGAC 1915  
Qy 1440 ggggagccagcagcttgcaaaatagatgtgacatgggattcccttgacccgcggt 1499  
Db 1916 GGGCGACCCACGCACTTGCATAATGAGATGACATGGAGATTCCCTGACTGCGCCCGGT 1975  
Qy 1500 tggagactacaacagcgttgaagtaggagagcctgaaatctatcgcgcagccttctgtg 1559  
Db 1976 TGGGACTACACACGCGTGAAGTAGGAGAGACTTGAANAATCTATGCGCAGGCTTGGTG 2035  
Qy 1560 gcgggtctccgggagcgtcgaagcgcgcacataatctgtgctaaggtaagaaggtgtg 1619  
Db 2036 GCGGGTCTCCGGGGCGCTCAAGAGCGCCACTAATTGGCTAAGTAAGAGAGGTGATG 2095  
Qy 1620 cagggacccgaagaacccctccgtatcttcttgagagctcgaatgaaagccttcagcgg 1679  
Db 2096 CAGGAGCCGAACGAACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTACGGGG 2155  
Qy 1680 ttcaccccttctgacccctcaccctcagagcccaagaacccctcagtgcccttgccctatt 1739  
Db 2156 TTCACCCCTTTGATCTTACCTCGSAGGCCACGAAGGCTCAGTGCCCTTGCTTCTT 2215  
Qy 1740 gggcagctcggtccttgatatacagaagaactcagagactcgaaggtttacagaggtc 1799  
Db 2216 GGGCAGCTGGCTCTGATATATCAGAAAGACTTCAGAGACTGGAAGGTTACAGAGGCT 2275  
Qy 1800 gagttacgtgatactagtagagagcagaagaagtgatattacagaagggagagaagaag 1859  
Db 2276 GAGTTACGTGATCTAGTAGAGAGCAGAGAAAGGTGATTACAGAAAGGAGACAGAAAG 2335  
Qy 1860 gagaaggaacagagaagaagaagagagaagaagaagggagagaagcgtgatagagcg 1919  
Db 2336 GAGAAGGAACAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2395  
Qy 1920 caagaagaacttctgactaagaatcttgcgcagctgtgttgaagggaagaagcagaagga 1979  
Db 2396 CAAGAGAAAAATTTGACTTAAGATCTTGCCCGAGTGTGAAGGAGAGAGAGAGAGAG 2455  
Qy 1980 agagagagagattttagaanaattagtcagagccctagaacgtcaggaacccgaggcaat 2039  
Db 2456 AGAGAGAGAGATTTTAGGAAAAATTAGTCAGGCCCTTACAGCTCAGGAGAACCTGGCAAT 2515  
Qy 2040 aggagccacactcagacaagaagcagctgtgctfatgttaagaanaagaacacttggcaga 2099  
Db 2516 AGAGCCCACTCGACAAAGAGACAGTGTCTGATTTGTAAGAAAAAGGACACTGGGCAAGG 2575  
Qy 2100 aactgcaccaagaagaagaacagaagcagaagtlcctagatcttagaagaagaataaat 2159  
Db 2576 AACTGCCCAAGAGAGAAACAAAGAGACGAAGGCTCTAGCTTAGAAGAGATTAAGAT 2635  
Qy 2160 tagggagagcgggttcggacccctcccgagcccaaggttaacttgaaggtgagggg 2219  
|||||

Db 2636 TAGGGAGAGCGGGGTTGGACCCCTCCCGAGCCAGGCTTACCTTTGAAGTGGAGGG 2695  
Qy 2220 caaccagctgagcttccgtgtgataccgagcgaagatcagtgctgtacacaacta 2279  
Db 2696 CAACCACTGATGTTCCGTGTTGATACGGAGCGGAGCATTCAGTGTCTGACAAACATTA 2755  
Qy 2280 ggaactaaagaanaaaatccctgggtgagtgggtgcacagggcaagcgaatcca 2339  
Db 2756 GGAANAACATAAAGAAAAAATCTGGGTATGGTCCACAGGCGAACCGCATATCCA 2815  
Qy 2340 tggactacccgaagaacgcttggacttggaggtgggaaggtaaaccaactgttttcgtgc 2399  
Db 2816 TGGACTACCCGAGAACCGTTGACTTGGCAGTGGAGGAGGATACCACTGTTCTGATC 2875  
Qy 2400 atccctgagtgccacagtaacccttctagtagagacttaactgaaccaagatggagagtc 2459  
Db 2876 ATCCCTAGTGGCCAGTATCCCTTCTTAGTATGAGACTTACTGACCAAGATGGAGCTCA 2935  
Qy 2460 attccttctgaacagaagaagaagaagtgctgtgtgataaacaacccactgtgtg 2519  
Db 2936 ATTCTTTTGAACAGAGAGAACAGAGAGTGTGTGAATTAACAACCCATCACTGTGTG 2995  
Qy 2520 accctcaattagatgtagaatacgcataatttccccaagttaaagcctgtatcaagt 2579  
Db 2996 ACCCTCAATTAGATGATGATATGATGATATATTCCTCCCAAGTAAAGCTGATCAAGAT 3055  
Qy 2580 atacagctccgtgttgagagcagttccccaagcctgggcaagaacgcaaggtgtgtg 2639  
Db 3056 ATACAGCTGTGTTGGAGCAGTTTCCCAAGCCTGGGAGAAACCCGAGGAGATGGGTTTG 3115  
Qy 2640 gcaaaagaagtlcccccacagttatatacactgaagcagtgctacacagatcagtc 2699  
Db 3116 GCAAGAGCAAGTTCCCCACAGGTTATTAATCACTGAAGGCCAGTCTACACAGATACAGTC 3175  
Qy 2700 agaaagtaacccttggatgaagagcgtcgcgaagaagatltggcgcagatgttcaaatga 2759  
Db 3176 AGACAGTACCCCTTGAATGATGAGAGGCTCGAAGAGAAATTTGGCGGATGTTCAAAATTA 3235  
Qy 2760 atccaagagacatccctagcttccgtgtccatcccttggatactccctgcacagctt 2819  
Db 3236 ATCCAAAGGAGCATCCAGTTCCTGTGCAATCCCTTGGAAATCTCCCTGCTACCGGTT 3295  
Qy 2820 aggaagccttggagccaatgattatcgaaccagtaacagacttgaagaggtlcaataaag 2879  
Db 3296 AGGAGGCTGGGACCAATGATTAATGACACAGTACAGACTTGAGAGAGGTCATTAAGAAG 3355  
Qy 2880 gtgcagagactaaccccaaggtcccgaaacctataacctcttgagcgccctccgct 2939  
Db 3356 GTGCAGAGCATACACCAAGGTCGGAACCCCTTATTAACCTCTTGAGCGCCCTCCGCT 3415  
Qy 2940 gaacgagacttgtlacacagatcttgaacttaagaatgaccttctgccttgaatatac 2999  
Db 3416 GAAGGAGACTGTATACAGATTTGAGACTTAAAGATGCCCTTCTTGCTGCTGAGATTACAC 3475  
Qy 3000 cccactagccaaccacttcttgccttgaatggagagatccaggtlacgggaagaacggg 3059  
Db 3476 CCCACTAGCAACCGGCTTTTGCTTCGAATGAGAGATCCAGGTCGGGAAGAACCGGG 3535  
Qy 3060 cagctcaactcgtgaacccgacttgcgccaaaggttcaagaactcccgacaactcttgaaga 3119  
Db 3536 CAGCTCACCGGAGACCGGACGCCCCAAGGGTTCCAAAGCTCCCGGACCATCTTTGACGA 3595  
Qy 3120 ggcctaacagagagccttgcgcaactcagatattcaaacacccctcaaggtagccctcccgag 3179  
Db 3596 GCCCTACACAGGAGACTGGCCAACTTCAGGATCCAAACCCCTAGGTGACCTCTCCGAG 3655  
Qy 3180 taagtgtatgacctgtcttgcgagagcacaacaaacagagactgtctagaaggtacgaag 3239  
Db 3656 TAGGTGATGATCCTGCTTGCGGGAGGACCAACAAAGAGACTGTAGAAAGGTACGAAG 3715  
Qy 3240 gcaactactgtgaatgtctgacctaggtacagagcctctgtctagaagaagccagatt 3299  
Db 3716 GCACACTGCTGGAATTTGCTGACCTAGCTACAGAGCCTGTGCTTAAGAGGCCAGATT 3775  
|||||

QY 3300 tgcagagagagatgaataactcttggtgtacagtttgcgggacggacgagctgaacg 3359  
|||||  
Db 3776 TGcAGGAGAGAGGTAAcATACTTGGGTACAGTTGGGGGGCGGCGACCATGGCTGACG 3835  
QY 3360 gagcagcaggaagaaactgtatgtccagataccggcccccaaccacgccaagaatgaga 3419  
3836 GAGGcAGGAGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCCAAACAAGTGA 3895  
QY 3420 gagtgtttggagacgtgtatgttgagactgtgatacccggtgttgagacctaga 3479  
|||||  
Db 3886 GAGTtttGGGAGcAGTGGAtTTTGGACGTGGATCCCGGGGTTTGAGCCTTAGCA 3955  
QY 3480 gccccactctacccgcttaaccaaagaaagggagattctccctgggctctgaacacag 3539  
3956 GCCCAGCTTACCCCGTAAcCAAAAGAAAGGGGAATTCTCCGCGCTCTGAGACACAG 4015  
QY 3540 aagcatttgaatctatcaaaaaagccctgtcgaagcactgtctgtgccccctgtac 3599  
|||||  
Db 4016 AAGcATTTGATGCTATCAAAAAGGCCCTGAGCGCACCTGCTGCGCCCTCCCGAC 4075  
QY 3600 gtaactaaccccttacccttattgtgatgagcgtgaaggagtagcccgaggaattta 3659  
|||||  
Db 4076 GTACTTAACCCCTTACCCCTTATGTGATGAGCGTAAAGGAGTAGCCGAGGAGTTTA 4135  
QY 3660 acccaaacccctaggaacatgagagagacactgttgcctactgtcaagaagacttgact 3719  
4136 ACCCAACCCCTAGGACCATGAGAGAGACCTGTGCGCTACCTGTCAAAAGAGCTTGATCC 4195  
QY 3720 gtgagcaggtgtgtggcccgatgtctgaagcgtatcgacgtgtgagccatctgtcaag 3779  
4196 GTAGCCAGGTGTGGCCCGTATGCTGAAGGCTATGCGACCTGTGCCACTACTGCTCAAG 4255  
QY 3780 gacgttgaacaattgactttggagacagaatataactgttaatagcccccatgatatggag 3839  
4256 GAGGCTGACAAATTGATTTGGGAGACAGAAATATACTGTATAGCCCCCATGTGATTGGAG 4315  
Db 3840 aacatcgttcgcgaagccccagacagatgatacgaacgccccgaatgaaccactatca 3899  
4316 AACATCGTTGCGAGGCCCGACCGCATGTGATGACCAACGCCCGCATGACCCACTATCA 4375  
QY 3900 agcctgtcttcacagagagaggttcaactttgtcccaacagccgcttcaacctgtccact 3959  
4376 AGCGTGCTTCTCAAGAGAGGCTACTTGTGCTCCACAGCCGCTCTCAACCCCTGCCACT 4435  
QY 3960 ctctgcctgaagagactgtgaacagtgactcatgattggcatcaactattgattggag 4019  
4436 CTTCGTGCTGAAGAGACTGATGAACCAAGTACTCATGATTGCCATTAATTGATTGAG 4495  
Db 4020 ggaacttgggttcgcgaagaccccttacaagataccgctgaacttggagaagtgtaaccttg 4079  
4496 GAACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGATCGAGAGAGTGTAACTG 4555  
QY 4080 ttcaatgaacggaagcagctatgtgtggaaggtgaagaatgtgtggtggcgagtggtg 4139  
4556 TTCACTGACGGAAGcAGCTATGTGTGTGAAGGTAAAGAGATGGCTGGCGGTGGG 4615  
QY 4140 gacggagaccgcacgaatctgtggcgaagcctgtccgaaaggaactcagcgcaaaagct 4199  
4616 GACGGGACCCACACAGATGTGTGGCCACAGCTGTCCGGAAGGAATCTCACGCAAAAAGCT 4675  
QY 4200 gagctaatggcctcaacgaagcttggggtgtggccgaaagggaatcatalaactllat 4259  
4676 GAGCTATGGCCCTCAcGGAAGGCTTTGGCGCTGGCCGAAGGGAATCCATTAACATTTAT 4735  
QY 4260 acggacagcaggtatgcttlttgactgacacgttacaaggggccaatcataaagaag 4319  
4736 ACcGGAAGcAGATGTACTTTGGACTGCAcGTACAGGGGGCCATCTATTAACAAAG 4795  
QY 4320 ggggttctaactcagcagggaggaataaagaacaagaaggaatcttaagcctata 4379  
4796 GGGTGTCTTACCTCAcCAGGAGGGAATATAGGAACAAAGAGGAATTTCTTAAGCCTATTA 4855  
QY 4380 gaaaccttaacttgcacaaagagctagctattatlaactgtcctggacatagaagcc 4439  
4856 GAAGCTTACATTTGGCAAAAAGGCTAGCTATTATATACTGTCTCGACATAGAAAGCC 4915  
QY 4440 aaagatcataatcagaaggaacagatgtgtgacgggtgtgccaagagcagccag 4499  
4916 AAAGATCTCATATCTAGAGGGAACcAGATGGCTGACCGGGTTGCCAAGCAGCACCCAG 4975  
QY 4500 gctgttaacctctgccttaataagaacgccccaaagccccgaacccaagcagatgac 4559  
4976 GCTGTTAACCTTCTGCTTATATAGAAACCGCCCAAGGCCCAAGCCGACAGACAGATAC 5035  
QY 4560 accctagaagacttgcgaagaagataaaagaatagaaccagttcttgaagctccgaagg 4619  
5036 ACCCTGAAGACTGGCAAGAGATMAAAAGATAGACACTTCTTGAGACTCCGAGGGG 5095  
QY 4620 accgtctaactcataatggaaggaatcctgtcccaacaaagaggtttagaatalgtc 4679  
5096 ACCTGTATACCTCATATGTGGAAGGAATCTGCCCAACAAAGAGGGTTAGATATGTGC 5155  
QY 4680 caacagatcatcgtctaaacccaactagaactaaacacttgcagcagttgtgaaga 4739  
5156 CAACAGATATCATGCTCTCAcCCACCTAGGACTTAAACCTGTGACAGTGTGTGAGAAcA 5215  
QY 4740 tccocctlatcatgtctgtgagctacagagtggtgtgactgtgtgtcaaacattgtg 4799  
5216 TCCCTTATCATGTTCTAGAGCTTACCAGAGAGTGGCTGACTCGTGCTGAACATTGTGTG 5275  
QY 4800 cccctgcagactgtgttaatgtctaatcccttcagaataactcacaagaagaagactaaagg 4859  
5276 CCTCGACAGCTGTTAAATGCTTAATCTTCAAGATACCTCCAGAAAGAGACTAAGGGGA 5335  
QY 4860 agcccccagggcgtcactcaggggaagtggaacttactgaaggttaaaacggcgttaaaag 4919  
5336 ACCCAcCAAGGCGCTACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTTAATAACGA 5395  
QY 4920 aacaaatataatgtgttltgttagacacacttcaagaatggttagaggtatctact 4979  
5396 AACAAATATCTATTGTGTTTGTAGACACCTTTTCAAGATGGGTAGAGGCTTATCTACT 5455  
QY 4980 aagaaagagacttcaacggtgtgtgtgaagaataacttgaaggaatlttccaaagatt 5039  
5456 AAGAAAGAGACTTCAACCGTGTGTAAAAAAATATCGAGGAATTTTCCGAGATT 5515  
QY 5040 ggaatacctaaggttaaggtgcagacaaatgttgcagtttcgttggcccaagtaagtag 5099  
5516 GGAATATCTTAAGTATGGGTCAAGACATGGTCCAGCTTTTGTGTCACAGGTAAAGTGA 5575  
QY 5100 ggaactggccaagataatgggattgaattgaaactgtcatgtgatalacagaacccaagc 5159  
5576 GGACTGGCCAGATATTGGGGATTGATTGSAACGTGCATTGTGCATACAGACCCCAAGC 5635  
QY 5160 tcaagaacaggttagagaggaatgaatagaaccataaagaagcccttacaacttgaacaa 5219  
5636 TCAGGACAGGTAAAGAGAGATTAATAAACCAATTAAAGAGACCCCTTACCAAAATTGACACA 5695  
QY 5220 ggaactggcatlaatgaattgatatgtctcctgccttggcttgaattgaagtgaaagac 5279  
5696 GAGACTGGCATTAATGATGTGATAGCTCTCTGCCCCCTTGTGTGCTTTTAAGGTTAAGGAAC 5755  
QY 5280 acccctgacagtttggctgacccctatgaatgtctctacagggggaacccccccgttg 5339  
5756 ACCCTGAGACAGTTTGGCGTACCCCTATGAATTGCTACGGGGGAGCCCCCGGTG 5815  
QY 5340 gcaagaatgtccttgaacataagtgctgaatgtctgtcttcccaagccttgttctcagg 5399  
5816 GTAGAAATTTGCTTCTGATCATATGTGATGTGCTGCTTCCACGCTCTGTCTTCTAAGG 5875  
QY 5400 ctcaaggtcgtcaggtgtgagcagggacgttgggaagcaggtccggggagggcctacta 5459  
5876 CTCAAGGCGCTCGAGTGGGTGAGGCAAGCCTGTGAAGCAGCTCCGGAGGCTACTCA 5935  
QY 5460 gg---aggaagacttgcagatgccacatgcttccaaagtgtgagattcagttcattag 5516



Db 5936 GGAGAGAGACTTGCAGATGCCATCGCTCCAGCTGGAGATGCTATGTTAGA 5995  
OY 5517 cgcacccgtgagaagaacctcgaagactcgttgaagaagactatctcfracttbbac 5576  
Db 5996 GCGCACCGTGGAGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTGACTTTTGACC 6055  
OY 5577 acaccacgctgtgaagtcgaagaatcccccacttgaatccatgatcccaagtaag 5636  
Db 6056 ACACCAAGGCTGTGAAGTGAAGGAAATCTCCACCTGATGATGATGATCCAGTTAG 6115  
OY 5637 ccgagcgacactcccgactcgggtgtgaagccgaaaaagactgaatccccctaaagt 5696  
Db 6116 CTGGCGCACCTCCGACTCGGGGTGAGAGCCGAAAGACTGAAATCCCTTAACTT 6175  
OY 5697 cgcctccatcgccgtgttcttacttcaacataactccccagcgagtagtaaacgct 5756  
Db 6176 CGCCTCCATCGCCTGGTTCCTTACTTACCAATTAACCTCCAGCGCGATGAACGCT 6235  
OY 5757 tatagaagctcgaaaccccaatagaaccttacccttacttggctgaatltgaacctga 5816  
Db 6236 TATAGACAGCTCGAACCCTATAGACCTTATCCCTTACCTGGCTGATTTATGACCTGA 6295  
OY 5817 taaggtgtcaactgaatgaagcactcgaagtgttgcctcctagaagcgacccgtgtgacctga 5876  
Db 6296 TACGGGTGTCACGTGAATAGCACTCGAGGTGTGCTCTAGAGCGCACTGTTGGCTGA 6355  
OY 5877 actgcatcttgcctccgatgtgattaaaccccgctgttaagaacacacactcccaactagt 5936  
Db 6356 ACTGCATTTTCGCTCCGATTTGATTAAACCCGCTGTTAAAGCACACACTCCCAACCTAGT 6415  
OY 5937 ccgtagtatagggttctatgtctgcagagcaacagagaagaataactgttgggttc 5996  
Db 6416 CCGTAGTTATGGGTTCTATTTGCTGCCCCAGCACAGAAAGAAATACGTGGGGTTTC 6475  
OY 5997 ttgggaatccctctgtgaaggaatgaagctgtcaccctccaagatgaagaacggaatg 6056  
Db 6476 TGAGGAATCCTTCTGTGAGGAAATGAGCTGCTGCTCAACGATGGAACCTGGAATG 6335  
OY 6057 gccatctctctccagagaccggtlaaatactccttctgtaacttccgcccggcgcaagta 6116  
Db 6536 GCCATCTCTCTCCAGSAGCCGGGTAATTTCTCTTGTCAATTCGCGCCGGGCAAGTA 6595  
OY 6117 caaataatgaataataataagaatgaagctgtcccccactcgaactgaatatactaa 6176  
Db 6596 CAAAGTATGAACTATATATAAGTAAAGTAAAGCTGCTCCCACTGACACTTAAATTAATCTAA 6655  
OY 6177 gataagttcaactg -aaaggaacacagaaataattcaaaagtgaataagtgtatga 6234  
Db 6656 GATAAGTTTCACTGAATAAAGAAACGAAACGAAATATTCAAAAGTGATTAATGTATGAG 6715  
OY 6235 ctgggaatagtttttat -tataggcggggaagcagggtlcaactttaaactcgcct 6293  
Db 6716 CTGGGAAATAGTTTATTAATAATATGCGGGGAGGAGGAGGTCACATTAACTTTCGCT 6775  
OY 6294 taggaatgaagaaggggacaaacccctctgtgcaatgagggcccgataaagtactgtcga 6553  
Db 6776 TAGATATGAGACGGGACAAACCCCTGTGGAAGTGGAGCCGATTAATACTTGTGCTGA 6635  
OY 6354 acaagggcccccggcccttgagcaacccgcaaaacttgcggttgcctcaatlaactcgcct 6413  
Db 6836 ACAGGGCCCGCCGCTTGAGAGCAACCGCATTAATTGCGGTCGCCCAATTAACTCCTGCT 6895  
OY 6414 ggcgctgacataacacagccgcttagcaacgtaaccatgtatgalttccataccaac 6473  
Db 6896 GCGGCTTGACATTAACAGCGCGCTTAGCAACGATGACACGATGATTTCTACCAACAC 6955  
OY 6474 ggcctagaactccccaggttctctgttgaagacagacagagacttgaatctatccaa 6533  
Db 6956 GCTTAGAAACCTCCCAAGTGTCTGTGTAAGACAGGACAGACTCTTCACTGCTCAATCA 7015  
OY 6534 gggaagttccaagcctcaactcaacggaacccctgagtcacttctctgttgccttg 6593  
Db 7016 GGGAGCTTTCAGAGCCATCACTCCAGCCGATGATGCCACTTCTTGTGGCTTGG 7075  
OY 6594 tctatcccaaggccctccttattatgaagggaatgaagaagaagaataatcaatgtac 6653  
Db 7076 TCTATCTTAAGGGCTCCTTATTATGAGGGATGGCTTAAGAAAGAAATTAATGTCAC 7135  
OY 6654 caaagacatagaatacaatgaatgaatgggttcccgaaataagacttaccctcaatgaat 6713  
Db 7136 CAAGAGACTAGAAATCAATGATGATGAGGGGTCGGAATAATGACTTACCTCAGTAAGT 7195  
OY 6714 ttccgggaagggaactgatatagaagaagctcccccattcccaacacacttctgtatag 6773  
Db 7196 TTCCGGGAAGGGACATGATGATGAGAAAGCTCCCATTCACCAACACTTGTGTATAG 7255  
OY 6774 taactgtgttatagaacagccctcgaataatgaatattagtaactgtgtataaagtg 6833  
Db 7256 TACTGTGTTTATGACAGAGCCTCAGAAATAGATTTAGTACCTGGTTAAACAGGTG 7315  
OY 6834 gtgggacatgaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 6893  
Db 7316 GTGGGATGCAATACTGGGTTAAACCCCTGTGTTCTCTCAGTCTTCAACCAATCCAA 7375  
OY 6894 agattctgttcaatgttccaaatgctcccgagtgatgaatgaatgaatgaatgaatgaat 6953  
Db 7376 AGATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7435  
OY 6954 ccttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 7013  
Db 7436 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7495  
OY 7014 agctgtatgtcgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 7073  
Db 7496 AGCTGTAATGCTGCGATTAGGAGCGCGCTGCTGAGGAACGAGGACAGCTGCCCTAT 7555  
OY 7074 caagagacacagcagcclagaagaagacttggtagcttaacttgcgcatacagagaaga 7133  
Db 7556 CACAGGACACAGCACTGAGCAAAAGACTTGTGATGACTGAGGACGAGGACGAGGACGAGG 7615  
OY 7134 tctccagccttaagaagagcttcttgaacacactgaagaagctcctgaacttcttctga 7193  
Db 7616 TCTCCAGGCTTGAAGAGGATCTGTTAGCAACTGAAGAGGCTCCGACTTCTTGTCTGA 7675  
OY 7194 agtgtgttcaagaagcaggaaggaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 7253  
Db 7676 AGTGTCTTACGAACCGGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 7735  
OY 7254 tgcagccttaagaagaagtgttcttctatgaatgaatgaatgaatgaatgaatgaatgaatgaat 7313  
Db 7736 TGCAGGCTTAAAGAAAGATGTTGCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 7795  
OY 7314 catgaacagcttagaagaagaatgaagaagcgtcgaagaagaagaagcgtcgaagaag 7373  
Db 7796 CATGAGCAAGCTTGAAGAAAGGTTAGAGAGGCTGGAAGGGAAGAGGCTGACCAAGG 7855  
OY 7374 gtggttgaagaatgttcaagaagcttccctgtgaatgaatgaatgaatgaatgaatgaatgaat 7433  
Db 7856 GTGTTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7493  
OY 7434 ggggcccccttagtgccttcttcttacttcaatgttggccttgaatgaatgaatgaat 7493  
Db 7916 GGGACCCCTAGTAGTCTGCTCTCTGTTACTTACATGTTGGGCTTGAATTAATAGTT 7975  
OY 7494 tgttgccttcttgaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 7553  
Db 7976 TGTGCTTGTGTTAGAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8035  
OY 7554 ccaagagccttgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 7613  
Db 8036 CCAAGGCTTCTGAGCCCAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8095  
OY 7614 ctatttaacagaagaagaatgtgggaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 7673  
Db 8096 CTATTAAACAGCAAGAGTGGGAAATGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 8155



```

- QY 7674 acccgaagtaataaaagctctaaatgccccgaattacaagccctgctgctgca 7733
      |||||||
Db 8156 ACCCAGGAAGTAATAAAAGCTTAATAAGCCCCGAAATTCAGACCCCTGCTGGTGC 8215
      |||||||
QY 7734 gtaaatagtaagaagtaacactctctatgttccaggcctgctatctggccta 7733
      |||||||
Db 8216 gtaaatagtaagaagtaacactctctatgttccaggcctgctatctggccta 8215
      |||||||
QY 7794 agatacagaagtaagtaagtaacactctctatgttccaggcctgctatctggccta 7733
      |||||||
Db 8276 AGATAACAGGAATGAGTTGAGTATCGCTTATCTGCGATTCTGTAAGAACTGACGCGCAC 8335
      |||||||
QY 7854 atagaagaatgtatcaacatgacagccctagtgacctatctcaactgcaatctgca 7913
      |||||||
Db 8336 ATAGAAGAAATGATTAACATGACAGCCCTAGTACCTATCTCACTGCAATCTGTCTAC 8395
      |||||||
QY 7914 tctgcccagagagccagccagatgagagccctggagactatcttaaatgttgctcac 7973
      |||||||
Db 8396 TCTGCCAGAGAGCCAGCCAGATGCGAGCTCCGAGCTATTTAAATGATGTGCTCAC 8455
      |||||||
QY 7974 ggaagcggagcctcgcataatcttaaaatgattgctcagagcggagcctcgcataat 8033
      |||||||
Db 8456 GGAGCGGGGCGCTCGATATTTTAAATGATTGGTCCAGCGAGCGGCGCTCTCGATATT 8515
      |||||||
QY 8034 ttaaaatgattggttltgacgcagagccttltgtgaacccataaaagctgtccga 8093
      |||||||
Db 8516 TTAATATGATTGTTGTGACGACAGGCTTTGTTGAACCCCATTAAGCTGTCCGGA 8575
      |||||||
QY 8094 ttcgcactcggggcgagcagcctcctcagcctgctgtgtgacagcgtgtggcccgagc 8153
      |||||||
Db 8576 TTCCGCACTCGGGGCGGAGCTCTTACCCCTGCTGTGTGACAGCTGTGGCCCCAGCG 8635
      |||||||
QY 8154 cgcctggaataaaatcctctgtctgttgcataaaa 8191
      |||||||
Db 8636 CGCTTGAATAAAATCCTCTTGTGCTTGTGATCAAGA 8673
      |||||||

RESULT 5
LOCUS PEN133816 8918 bp DNA VRL 07-DEC-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 33).
ACCESSION AF133816
VERSION AF133816.1 GI:6688945
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE Porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus.
REFERENCE 1. (bases 1 to 8918)
AUTHORS Caendera,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2. (bases 1 to 8918)
AUTHORS Toenjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
source
location/Qualifiers
1..8918
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"
/db_xref="taxon:61673"
/clone="33"
/note="grown in PK15 (ATCC No. CCL-33) cell line"
LTR
misc_feature 545

/note="cap site"
710..727
/note="g1y4"
1154..2725
/gene="gag"
1154..2725
/gene="gag"
/function="retroviral capsid"
/product="Gag protein"
join(1154..2725,2729..6313,6189..8162)
/note="viral polypeptide"
/codon_start=1
/protein_id="CAB65339.1"
/db_xref="GI:6688946"
/translation="MGQVTVPLSLTLDHWTVRSAHNLISVQKKGPWQFCASEMP
TRDVGPSSEGTNSETILLAVKAIIFQGTSSHPDEPYLLTQDLAEDPPVAKTFLN
KRKPGRIILALGSEKNSHAERVSERPRIYPIIEPPVPEQPPVPPVAGAVRG
PSAPGAPVVEGPAAGTISRSGATPERDEIALLPLRTYGPMPGOLQLOLPYFSS
ADLYNMKTNHPPESEDPORLTGLVESLSHOPMDCCOOLLQTLFTTERERITLLA
RKNVPGADGRPOLONEIDMGPRPLRPGMDVYTAGRESIKYIRQALVAGLRGASRP
ITLAKRREVMQGNPEPSEFLEIRIMARFRPTDPPTSAQASVALFVIGSALDIR
KILQVLEGLQDAELDLVLEAKVYIRRETEKQREKEREREKREKREKREKREK
KILAAVEEKSSEERERDRKIRSGPROSGNINRTPLDKQCAVCKRGMARNCPK
KNGKPKVLTLEDDKGRGSDPLEPRVTLKEGQPEFLVDGAEHSVLLQPIGLK
KEKSWMGATGOROYPMTTTRRTDLAVGRVTHSLVIECPVPLIGDILLTKMAOI
SPEQGRVSVNNKPIVYLTLOLDEYRLSPQKPPDODIOSMLDOPPOANAEATGK
LAKOVPRVIOUKASNTPVSVQVYPLSRARGTIPVHORIQQGILVAVPSPWNTPL
LIVRKGTINDYRPVODLRVNRVODIHETVNPYNLSALPERNWTVDLIDKAF
CLRHPTSQPLPAPFMRDPTGRTQLTWTPLRPOFKNSPTLEDEALHDLANERIOH
POVTLQYVDDLLAGATKODLEGTKALLLESLDGRASAKKQALORREVTYIGS
LNGGQWMLTEARKKTIVVLPAPTAKOYREFGTGFCRLMIPGFATLAAPLYPIKE
KEGFSWAPRQKAPDAIKKALLSAPALAPDTRKPTLYVDERKVASGVITQTLGPM
RBPVAVYSKKIDPVASGMPVCLKATAVAAILVKDADKTLTGONTIVVAPHAENVRQ
PPDRKNTNARMTHTOSLTERVTPAPRALPAPLLEEDTEPTVHCDOHLLEETG
VRRKDLIDPLTGEVLTWFTDSSVYVEKGRMAGAADVGTHTMASLREGTSQAKAE
LMALQALRIAEAGKSINITYDSRFATFAHGAIVYKORGLTSGREIKKEELISL
LEALILPKRLAIHCPGHOKAKDLISRGNOMADRAKOAQAVNLPIIETPKRPEP
ROYTLEDMOEIKKIDQFSEPTGCTYSTGKELLPKRGLGYVOOIHRLTIGRKHIO
OLVPTSPYHVLRLPGVADSPVHCYPCOLVANNANRIRTPGRLKRSHPGAMVEDPTE
VAPRAKYKKILVFDITSGWAEIPIKKESTVAAKILLEIPREFIPVVISDNG
PAFVAQVOSGLAKILGIDMKLHCAYRPOSSQOVERMNTIKETLKTLETGINDMIA
LLPEVLFRNTPGOGLTPEYELAYRPGPVEIVLASVADVLSQPLFSLTKALEVY
RORANKOLREAVSGEDLOVHRFOVGSVYVRRRAGNLEFRMGKPYLVLTLPVAV
KYEGISTWHTSHVKLAPDPDGSQWAEETEMPKIRLRIYVYNNNSPGOIHPTLS
RHLPRGSEPKRIRPLPSFASLMTPLITITPOASKKLIDSSNHRPLSLTFLITDP
DTGVTVNSTRGVAPRGITWPELHFCRLRLINPAVASTPNNLVRSYGFCYCCPTEKRYC
GDGSEFCRMSVCVINDGDKMWPISLIDRVKFSVNGSPKRYVMKLYKXKSCSPSD
LDLYKISTTEKGKOBENIKWINGMSGIVFYKYGGAGSTLIRLRETEGPEVAVG
FDKVLAEQGPALPEPNNLPVOLTSLCPDLTOSPSNGTGLIPLNPRNSGVPVKT
GORLSLIOGAFOALINSDPDATSCWCLSSGPPYVGMKAKEKENVTKRHRQNCW
GSRNKLITLTVSGKGTCTGKAPRSHOHCISTVYEQASMOYLVPGYNNRMAKNTGL
TPCVSTSVFNOSKDPVWQVQIPRYTHPEEVLDEVDYRYNRKREVSITLAVMLG
LGTAVGVGTGTAALITGPQOLEKIGLEHAAVDEYDRIYNRKREVSITLAVMLG
ONRGLIDLLFLREGICLALKKECCFYVDHGAIRDSKILRELERERREREDOGG
FEGMFRSPMMTTLTSLALGPLVILLVLLVGPCLINFEVAFVERYSAVGIMVLROQ
YGLISQGETDL"
2729..6313
/gene="pol"
/function="retroviral protease and polymerase"
/product="pol protein"
2729..6313
/gene="pol"
/gene="pol"
6189..8162
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6189..8162
/gene="env"
8212..8918
8811..8816
BASE COUNT 2327 a 2245 c 2324 g 2012 t
ORIGIN

```



QY 2100 aactgcccgaagaaggaaacaaaggaacgaaggtcctagctctagaagaagataaagat 2159  
DB 2666 AACTGCCCAAGAGGAGAAACAAAGGACCGAAGGTCCTAACTTAGAAGAAGATAAAGAT 2725  
QY 2160 taaggagagaggggttcggaacccctcccgagcccgaggttaacttgaagtgagagg 2219  
DB 2726 TAGGGGAGACGGGTTGGAGACCCCTCCGAGCCAGGGTAACTTTGAAGGTGAGAGGG 2785  
QY 2220 caacgaattgagttcctgattgataccgagcgagacatcagtgctgctacaacatta 2279  
DB 2786 CAACCAAGTGAAGTTCCTGTTGATACCGGAGCGAGCATTCAGTGTCTACAAACCTTA 2845  
QY 2280 ggaacactaaagaaaaaatacctgggtatggtgccaagggcaagcgagatcca 2339  
DB 2846 GGAATACTAAAGAAAAAATCCTGGGTATGGGTGCCACAGGCAAGCGATATCCA 2905  
QY 2340 tggactaccggaagaacccgttgacttggagtgagcggttaaccaactgcttctgctc 2399  
DB 2906 TGGACTACCGGAAGAACCGTTGACTTGGCAGTGGGAGCGGTAAACCCACTGTTCTGCTC 2965  
QY 2400 atccctagtgcccgagtaacccctctagtgagagactactgacgaagatggagatcca 2459  
DB 2966 ATCCCTAGTGCCAGTACCCCTTCTAGTAGAGACTTACTGACCAAGATGGAGCTCAA 3025  
QY 2460 attcctttgaaacagaagacagaagtgctgtgaataacaacccatcaactgtgtg 2519  
DB 3026 ATTTCTTTGAAACAGSAGSAGAACGAAAGTGTGTGATTAACAACCCATCATGTGTTG 3085  
QY 2520 accctccaattagatgataatgataatgactatctccccaagtaaaagcctgaatcaat 2579  
DB 3086 ACCCTCCAATTAGATGATGATATGATATATCTCCCAAGTAAAGCCTGATCAAGAT 3145  
QY 2580 atacaagtcctggttgaggaaggttccccaagcctggagcaaaacgagagatgggttg 2639  
DB 3146 ATACAGTCCGTGTTGGAGAGCTTTCCCAAGCCTTGAGGCAAAACCCAGGAGTGGTTG 3205  
QY 2640 gcaaaagaaagttcccccacaggtatctcaactgaagcgagtgctacacagatcaagtc 2699  
DB 3206 GCAAAAGAAAGTTCCTCCACAGGTTATTAACCTGAAGGCCAGTGTCTACACAGTATCAGTC 3265  
QY 2700 aagcagtaaccccttgatgtagaagagctcgaagaagaattggcgcgaagtccaagatta 2759  
DB 3266 AGACAGTACCCCTTGATGATAGAGGCTCGAAGAGAAATTTGGCCGATGTTCAAAAGTTA 3325  
QY 2760 atccaaagggcatcctcagttcctgctcaatcccttggaataactccctgctacccggt 2819  
DB 3326 ATCCAAAGGGGCAATCTAGTCTGTGCAATCCCTTGGAATATCTCCCTGCTACCGGTT 3385  
QY 2820 aggaagccttggaaccaatgatalcgaacagtaacagacttgagagagltcaataaag 2879  
DB 3386 AGGAAGCCTGGGACCAATGATTAATGACCAAGTACAGGACTTGAGAGAGTCAATAAAG 3445  
QY 2880 gtgcagagataaaccacaaggtcccgaaaccttaacctctgagcgccctccgcct 2939  
DB 3446 GTGCAGAGATACACCAACGAGGTCGGAACCCCTTAATACCTCTTGAGCCCTCCGCTC 3505  
QY 2940 gaacggaactgtaacagatattggaactaaagatgcttcttcgctcgatgataac 2999  
DB 3506 GAACGGAACGTGATACACAGTATTGGACTTAANAAGATGCTTCTTGCTTGAGATTAC 3565  
QY 3000 cccactagcaaacacttttgcctcgaattgagagatccagatlcagtgaggaacacgg 3059  
DB 3566 CCCACTAGCAACCCGCTTTTGGCTTGAATGAGATCCAGATCCAGSAGAAACCGG 3625  
QY 3060 cagctacttgaccccgactgcccgaaggttcaagaactcccgacaactcttgaaga 3119  
DB 3626 CAGCTACTGTAACCCGACTGGCCCAAGGGTTCAAGAACTCCCAACATCTTTTGAACAA 3685  
QY 3120 ggcctacacaggaaccttgccaactcaagatccaaacccctcaggttaacccctccag 3179  
DB 3686 GGCCTTACACAGGAGACTGGCCCAACTTCAAGATCCAAACCCCTCAAGTGACCTCTCAG 3745  
QY 3180 taagtgatgacctcttctgctggagagccaaacagagactgcttagaagtgacgaag 3239  
DB 3746 TACCTGATGAGACTGCTTCTGGGGGAGGACCAACCAAGACTGCTTAAGAGTACGAAG 3805  
QY 3240 gcaactgctgtaactgtctgaactaggctacaagaacctctgctaaagaagccagat 3299  
DB 3806 GCACCTACTGTAATTTGCTGACCTAGGCTACAGAGCTCTGTGCTTAAGAAGCCAGATT 3865  
QY 3300 tgcagagagaggttaactacttgggtgacgtttgacgggacgggagagatgactgag 3359  
DB 3866 TGCAGAGAGAGGTAACTATCTTGGGGTACAGTTTGGGGGGGAGGAGATGGCTGAG 3925  
QY 3360 gaagcaggaagaanaactgtagtccagataccggcccacaacacagccaaacagtgaga 3419  
DB 3926 GAGCGACGGAAGAAACTATGATCCAGATACCGGCCCCACACACCAACCAAGTGA 3985  
QY 3420 gagttttgggaacagctgagatlttgacagacttgatcccggttttgacactta 3479  
DB 3986 GAGTTTGGGAGACGCTGATTTTGGACAGCTGTGATCCGGGGTTTGGACCTTAGCA 4045  
QY 3480 gccccactacccgttaaccaaagaaggggattctcctggtctcctgagcacaag 3539  
DB 4046 GCCCACTCTACCCGCTTAACCAAGAAAGGGGAATTTCTCTGGGCTCTGAGCACAG 4105  
QY 3540 aaggaatttgatgatacaaaaagccctgctgagcgcaacctgctctgagccctcctgac 3599  
DB 4106 AAGGCAATTTGATGCTATCAAAAAGCCCTGCTGAGCGCACCTGCTCTG6CCTCCCTGAC 4165  
QY 3600 gtaactaaaccccttaaccccttaagtgtgatacgtlaagggagtagcccgagattta 3659  
DB 4166 GTAACATAAACCTTTACCTTTATGTGATGAGTGAGCGTAAAGGGAGTACGCCAGATTTTA 4225  
QY 3660 acccaaacccctagagcaatggagagacctgtgctcactgtcaagaagcttgaacct 3719  
DB 4226 ACCCAAACCCCTTAGACCAAGAGGAGAGACCTGTGCTACCTGTCAAAAGAGCTGTATCT 4285  
QY 3720 gtaagcagtggttgcccgatgctgtaagaggtatcgaagcgtgtgagcaactgtgaag 3779  
DB 4286 GTAAGCAGTGTGGCCCGATGCTGTAAGGCTTATGCAAGCTGTGGCCATCTGCTCAG 4345  
QY 3780 gacgctgaacaaatgacttgggacagaataaactgtatagcccccatgcatltgag 3839  
DB 4346 GACGCTGACAAATGACTTTGGGACAGATATTAATGTAATAGCCCCCATGATTTGAG 4405  
QY 3840 aacatcgttcgcaagccccacagacagatgtagtgaacaaagcccgatgaacccatcaa 3899  
DB 4406 AACATCTGTTGGCAACCCCAAGCAGATGATGACCAAGCCCGATGACCCACTATCA 4465  
QY 3900 agcctgctcacaagagaggtacacttgcctcacaagccgctcaaacctgaacct 3959  
DB 4466 AGCCTGCTTCTACAGAGAGGCTCACTTCCGCTCCACAGCCGCTCTTAACCTGCACT 4525  
QY 3960 ctctcgtcgaagagactgtgaacagtgatcatgatagtgccatcaactatgtatgag 4019  
DB 4526 CTTCTGCTGAAGAGACTGATGAAACAGTACATGATGTTCCATCAACTTTGATTGAG 4585  
QY 4020 gagactgggtccggaagaccttaacagataacgcgtgactgtgagaaggtgtaacctg 4079  
DB 4586 GAGACTGGGTCGCAAGAGACCTTAACAGACTATCCGCTGAGTGAAGAGTCTAACTCG 4645  
QY 4080 ttaactgacgaaggaagtagtggtagaaggaaggaagtgctggcgagtgag 4139  
DB 4646 TTCACTGAGGAGAGAGCTATGTGTGTAAGATGAAGAGTGTGGGGGCGCGGTGAG 4705  
QY 4140 gaecggaccgcagatctgggcagagcctgcccgaagaagactcagcgcgaagact 4199  
DB 4706 GACGGACCCACAGATCTGGGACAGCACTGCGGAAGAAGACTTGAAGGCAAAAGGCT 4765  
QY 4200 gagctatgacctcaagcaagctttgagctggcggaaggaattcaataactat 4259  
DB 4766 GAGCTATGAGGCTTCACAGCAAGCTTTGCGGTGGCCCAAGGAATCAATTAACATTAT 4825  
QY 4260 acggaagcagtagtgccttgcgactgcaactgaaggggacatctaaacgaag 4319

Dh 4826 ACGGACAGCAGGTATGCTTGGGACTGCACACGTACACGGGGCCATCTATTAACAAGG 4885  
Qy 4320 ggggtgttacctcaagcaggaggaataaagaacaagaagaaattcctaagctatla 4379  
Dh 4886 GGGTGGCTTACCTCAGAGGAGGGAATAAAGACAAAGAGAAATTCTAACCTATTAT 4945  
Qy 4380 gaagccttaccttgcacaaaggctagctattatatacactgtccctgacatcagaagacc 4439  
Dh 4946 GAAGCCTTACATTTGCGCAAAAGAGCTAGCTATTATACACTGCTCTGACATCAGAAAGCC 5005  
Qy 4440 aaagatctcatctatagaagaaaccagatgctgacccgggttccaaagcaggagccag 4499  
Dh 5006 AAAGATCTCATATCTAAGAGGAAACCATGTGACCGGGTTGCCAAGACGGGACCCAG 5065  
Qy 4500 gctgttaacctctgcctataatagaacacgcccacaaagcccccagaacccagacagctac 4559  
Dh 5066 GCTGTAACTCTTGCTGCTATATAGAAAGCCCAAGGCCCAAGACCCAGACAGCAGTAC 5125  
Qy 4560 accctagaagaactgagcaagagataaagaatagaccagttctctctgagactccgagagg 4619  
Dh 5126 ACCCTAAGACACTGGCAAGAGATAAAAAGATAGACACAGTCTCTAGACTCCGGAGGG 5185  
Qy 4620 accgtgcatctacatataaggaaagaaatctgcccacaaagagggtagaataatgctc 4679  
Dh 5186 ACCGTCTATACCTCATATGGGAAGAAATCTGCCCCACAAAGAGGTTAGAAATATGTC 5245  
Qy 4680 caacagatatcgtctaacccacactagaagaactaaacactgacagagttgttcagaaaca 4739  
Dh 5246 CAACAGATATACATGCTCTCACCCACCTAGGAACCTAAGACCTGCAGCAGTGTGCAAGACA 5305  
Qy 4740 tccctcatcatctgttcgaaggtacacagagtggtgagctcagttgtaaaacttctg 4799  
Dh 5306 TCCCTTATATATGTTCTGAGAGCTACAGGAGTGGCTGACTGGTGGTCAAACTTGTGTG 5365  
Qy 4800 cccctgcaagcttgttlaatgtaactccttccagaatacctccagaagaagagactaaagg 4859  
Dh 5366 CCGTGCACAGCTGGTTATGCTAATCTTCCAGAAATACCTCAGAGAAAGACACTAAGGGGA 5425  
Qy 4860 agccacccagcgtctcaactgaggaagtgtgactctcagtgaggttaagccggtcctaatac 4919  
Dh 5426 AGCCACCCAGGCGCTCCTGGGAAGTGGACTCTAGAGGTAAAGCCGGCTAAATTCAGGA 5485  
Qy 4920 aacaatcatctattgtgttctgttagaacacttccagagtggttagaggttactact 4979  
Dh 5486 AACCAATATCTATTGGTTTGTGTAGACACCTTTCAAGATGGGTAGAGGCTTATCTACT 5545  
Qy 4980 aagaagaagactcaacccgtgtgtgtctaaagaanaatactggaaggaatttltccaagaatt 5039  
Dh 5546 AAGAAAGAGACTCAACCCGTGTGGCTAAAAAAATACTGAGGAATTTTCCGAGATT 5605  
Qy 5040 ggaatacctaagttaataaggtcagacaatgtgtccagcttctgttcccaagtaagtcag 5099  
Dh 5606 GGATATACCTTAAGTAACCGGGTCAGACAAATGTCTCAGCTTTTGTCCCGAGTAACTCAG 5665  
Qy 5100 ggaactggccaagatattggaggtattgtgaactcattgtgcatcagagccccaagc 5159  
Dh 5666 GGACTGGCCAAAGATATTGGGGATTGATTTGGAACCTGCATTGTCTATACACACCCCAAG 5725  
Qy 5160 tcaaggaaggttagaagagatgaatagaacatlaaagagaccccttaaccaattgacaca 5219  
Dh 5726 TCAGGACAGGTAGAGAGATGATAGAACATTAAGAGACCCTTACCAAAATGACACACA 5785  
Qy 5220 gaaactggcaatlaagatgtgactcctcctgcttctgttctttagaggttagaagaa 5279  
Dh 5786 GAAGCTGGCATTTATATGATTTGATAGCTCTCCCTCTCTTGTGTAGGGTTAGGAAC 5845  
Qy 5280 acccttgacagtttggctgaccccccataatgtaattgctcagggagagaccccccgttg 5339  
Dh 5846 ACCCTTGACAGTTTGGGCTGACCCCTATGAAATGTCTCTACGGGGACCCCCCGCTTG 5905  
Qy 5340 gcaagaatttgccttgcacatagtgctgagatgtgtgtcttccagccttgttctctag 5399  
Dh 5906 GTAGAAATGTCTCTGACATAGTGTGATGTGCTGCTTCCAGCCTCTGTCTCTAG 5965

Qy 5400 ctcaagcgcctcagtggttgaagcagcagctgtgaagcagctcccggaagcctactca 5459  
Dh 5966 CTCAGGGCGCTGATGGGTGAGGACACAGAGGTGGAAGCAGCTCCGGGAGGCGTACTCA 6025  
Qy 5460 gq---aggagacttgcaggttccacatcgccttccaagttgagatcagttcattgta 5516  
Dh 6026 GGAGAAAGGAGACTTCCAAAGTTCCACATGCTTCCAAAGGGGAGATTCAGTCTATGTAGA 6085  
Qy 5517 ggcacccgtgaggaagaaacctcgagactcgtgtgaagaggacctatctgacttctgacc 5576  
Dh 6086 CGCCACCGTGCAGGAAGAACTCGAGACTCGGTGGAAGGGCCCTTATCTGTACTTTGACC 6145  
Qy 5577 aacccaagcgtctgtgaagtcgaaggaatcccccactggaatccatgcaatccacgtlaag 5636  
Dh 6146 ACACCAAGCGGTGTGAAGTCAAGAAATCTCCACTGGATCCATCATCCACAGTTAAG 6205  
Qy 5637 ccggtgcacacccctccgactcgtggtgtgagagcccgaaagacttgaatcccttaagctc 5686  
Dh 6206 CTGGGGCCACTCCCGACTCGGGGTGAGAGCCGAAAGACTGAGAAATCCCTTAAGCTT 6265  
Qy 5697 cgcctccatcgcctgtgttccctactctaaacaaatacctccagcagtagtaaacgct 5756  
Dh 6266 CGCTTCATCGGCTGTTCTTCTTACTCTAACAATACTCCAGGCGCACTAGTAAAGCCT 6325  
Qy 5757 tatagaagcttgcgaaccccccataagaccttalccttaccctgtgtatattgacctga 5816  
Dh 6326 TATAGACAGCTTGAACCCCATAGAGACTTATCCCTTACCTGCGGATTTATGACCCTGA 6385  
Qy 5817 tacgggtgtcactgttaaaatagacactcgaaggtgtgtcctctgaagaccctgtgtgcctga 5876  
Dh 6386 TACGGGTGTACTGTAATAGACTCGAGCTCGAGGTCTCTCTAGAGGCACTGTGTGCTCTGA 6445  
Qy 5877 actgcatcttgcctcgcgattgtatlaaccgcgctgtltaaaagcacacctcccaactagt 5936  
Dh 6446 ACTGCATTTCTCTCCGATTGATTAAACCCGCTGTAAAGACACACTCCCAACCTAGT 6505  
Qy 5937 ccgtatgtatgtgttctatgtgtcgcagagcaagaagaagaatactgttgggtgtc 5996  
Dh 6506 CCGTATGTAATGGGTTCTATGTGCTGCCAGGCAACAGAAAGAAATCTGTGGGATTC 6555  
Qy 5997 tgggaatcctctgttaagagaatgagctgtcaccctccaaagatgagaatgaaatg 6056  
Dh 6556 TGGGGAATCTCTGTAGAGATGAGAGCTGCTACCTCCAAACATGAGATCGGAATG 6625  
Qy 6057 gccgatctctccaggaacccgggtaaatctcctctgttcaattccggccgggcaagta 6116  
Dh 6626 GCCGATCTCTCCAGGACCGGGTAAATTTCTCTTGTCAATTCCGGCCGGCAAGTA 6685  
Qy 6117 caaatgataagaactatataaagataaagctgtcctcccatccagacttaagtattctaa 6176  
Dh 6686 CAAAGTGATGAACATATATAAGATMAAGCTGCTCCCATCAGACTTATGATTTCTAAA 6745  
Qy 6177 gataagttcactg--aaaggaacacaggaataatlaaagaatgataatgataag 6234  
Dh 6746 GATTAAGTTCACTGAAAAAGGAAACAGAGAAATTTAAAAAGTGTGATTAATGATATGAG 6805  
Qy 6235 ctggggaaatgtttttat-tatagtgcggggagcaaggtgccacttaaccatctgct 6293  
Dh 6806 CTGGGGAATAGTTTATTAATATATGCGGGGAGACAGGTCACCTTAACCATTCGCT 6865  
Qy 6294 taggataagagacggggaacaaagacccctgtgtgcaatgagaccgataaagtaactgtcga 6353  
Dh 6866 TAGGATAGAGAGAGGAGGAAGAACCCCTGTGCAATGGGACCGATTAAGTACGTGCGA 6925  
Qy 6354 acaagggccccggccttgcagccacgcatlaacttgcggttgcacaaatlaactcgt 6413  
Dh 6926 ACAGGGGCCCCGGGCTTGAGGCCACCGCATTAACCTGCGGCTGCCAATTAACCTGCT 6965  
Qy 6414 gcggcctgacataacacacgcgctagcaacagtaaccaatgattgattcctacaaac 6473  
Dh 6986 GCGGCTGACATTAACACAGTACGCTGACAGCAAGTACACTGATGATTTCTTACCAACAC 7045

QY 6474 gctagaactccccaggtgttctgttaagacagacagagacttcaactca 6533  
 DB 7046 GCGTAGAACTCCCAAGGtTCTGTATAGACAGACAGACTTGTATCTCA 7105  
 QY 6534 gggagcttcccaagcatcaactccaccagacctgtatgcacttcttctgttgc 6593  
 DB 7106 GGGAGCTTCCAGGCATCACTCCACGACCCCTGATGACACTCTCTGTTGGCTTGG 7165  
 QY 6594 tctatcccaaggcctccttaltatgaggagatgctaaagaagaatcaatgtgac 6653  
 DB 7166 TCTATCTCTCAGGCTCTCTATATAGAGGAGTGGCTAAAGAAATTCATATGTGAC 7225  
 QY 6654 caaagagcatagaatcatatgacatgtgggtccgaaataagcttaccctacatga 6713  
 DB 7226 CAAGAGCATAGAAATCAATGATACATGGGGGTCCGAAATTAAGCTTACCTCAGAAAT 7285  
 QY 6714 ttccgggaaggagacatgacatagaaagaactcccccatcccaacacacttgcatalag 6773  
 DB 7286 TTCGGGGAAGGGGACATAGCATAGAAAGCTCCCATCCACCAACACTTGTGTATAG 7345  
 QY 6774 tactgtgttaltatgacagagcctcagaataacagatattagacctgtgtataacagtg 6833  
 DB 7346 TACTGTGTATATGACAGAGCCCTCAGAAATCAGTATTTAGTACCTGTTATTAACAGGTG 7405  
 QY 6834 gtggagatcaataactactgggtttaacccccctgtgttccacctgaacttcaacacacaa 6893  
 DB 7406 GTGGGATGCAATACTGGGTTTACCCCTGTGTTCACCTCAGTCTTAAACCAATCCAA 7465  
 QY 6894 agattctgttcatgtgtccaaatcgltcccccagtgtaactacactcctgagaaagtgtg 6953  
 DB 7466 AGATTTCGTGTCTGTATGTCCTCAATTCGTCCTCCGAGTGTACTACATTCGAGAAATGGT 7525  
 QY 6954 ccttgatgaataatgactatctgtgtataacagacaaagaagaacccgtatccctaccc 7013  
 DB 7526 CTTTGAATGATATGACTATCTGTATTAACCCGACCAAAAGAAAGCCGTATCCCTTACCT 7585  
 QY 7014 agctgaatgtcgtggaattagagacgagcgtgtgacgttagaagacagagacgtcctgtat 7073  
 DB 7586 AGCTGATATCTGTGATTTAGGAGCGGCGTGTGGCTGATAGACAGGAGACAGCTCCCTGAT 7645  
 QY 7074 caaagacacacagacagtagaagaagaacttggtgacatcatgctgagacatgacagaaga 7133  
 DB 7646 CACAGACACACAGCAGCTAGAGAAAGACTTGTGTAGTACATGCGGCATGACAGAGAA 7705  
 QY 7134 tctccagagccttaagagagcgtgttagcaacctagaagaagctccctacttctgtctga 7193  
 DB 7706 TCTCCAGCCTTGTAGAGAGTCTGTATGCAACCTTAGAAGAGTCCCTGACTTGTGTCTGA 7765  
 QY 7194 agtggtctacagaacacgagagagatagatctgtgtcttctaagaagaagtgtgtatg 7253  
 DB 7766 AGTGCTCTACAGAACCGAGGGGATTAAGTCTGTGTCTTCTTAAGAGAAAGTGTGTATG 7825  
 QY 7254 tgcagccttaaaagaagaatgtgtctctatgtatgactacacagagacatcagagactc 7313  
 DB 7826 TGCAGCCTTAAAGAAAGAAAGTGTCTGTCTGTATGATCACTCAGAGCCATCAGAGATC 7885  
 QY 7314 catgaagaacactagaaaaaaagttagagagcgtcgaagagaaagaagagcttcaacagg 7373  
 DB 7886 CATGAGCAACCTTAGAAGAAAGTGTAGAGGCGCTCGAAGGGAAGAGAGGCTTACCAAGG 7945  
 QY 7374 gtgtgtgaagagatgtgtcaacaggtctcctgtgatacaccacacttcttctgtctgac 7433  
 DB 7946 GTGTGTGAAGAGATGTTCAACAGGTCTCTTGTGATGACACCTCTCTCTCTGTGAC 8005  
 QY 7434 gggagccttagtagtctgtctctgttacttaccagttggccttgccttaataatggt 7493  
 DB 8006 GGGAGCCCTAGTATCTCTCTCTGTACTTACAGTTGGGCTTGTATTAATAGGT 8065  
 QY 7494 ttgtgcttctgttagaagaagatgagtagccagataatgttactttaggaagaaga 7553  
 DB 8066 TGTGTCTTGTTAGAAGACAGATGAGTGCAGTCCAGATATGTGTACTTGGCAACAGTA 8125  
 QY 7554 ccaagccttctgagccaagagaagaactgactccttagccttccagatttcaagatagaa 7613

DB 8126 CCAAGCGCTTCGAGCACAAGGAAGAACTGACCTTAGCCTCCAGTCTTAAGTATGAA 8185  
 QY 7614 ctattacaagaagaagtgtgggaatgaagaatgaataatgacactcaactccaga 7673  
 DB 8186 CTATTAAACAGACAAAGAGTGGGGAATGAAGAGATGAAGAAATGCAACCTTCCAGAG 8245  
 QY 7674 acccaggaagttaataaagaactcraaagtcgccccgaattacaacacttgcctgcca 7733  
 DB 8246 ACCCAGGAAGTAAATAAAGCTTAATATGCCCGGAAATCCAGACCTGCTGGCTGCA 8305  
 QY 7734 gtaaatagtagaagtgacacttctatltgtccagagcctgtatcctgtgcca 7793  
 DB 8306 GTAAATAGTAGAAGTACACTTCTCTATTTGTTCAGAGGCGCTGATCTGCGCTAAGTA 8365  
 QY 7794 agataacaggaatgagttgactaatgccttatctgtatctgttaaaactgacggacc 7853  
 DB 8366 AGATTAACAGGAATATGTTGACTAATCCCTATCTGATTCGTGTAATCACTGACGAC 8425  
 QY 7854 atagaagaattgattacacatgacagccttagtgacatcctaactgaactgtcac 7913  
 DB 8426 ATGAGAAGATTTGATTACACATTGACAGCCCTAGTACCTATCTCACTGCAATCTGTAC 8485  
 QY 7914 tctgccagagagccacagacatgcygacactccggagc-----tatt 7951  
 DB 8486 TCTGCCAGAGACCCACGAGATGCGGACCTCCGAGCTATTTTAAATGATGGTCCAC 8545  
 QY 7952 -----tatt 7955  
 DB 8546 GGAGCGCGGCTCTCGATATTTTAAATGATTTGTTCACAGAGCGCGGCTCTCGATATTT 8605  
 QY 7956 ttaaatatgtgtcacaagagcgggctcgtcatatttaaaatgattgtccatgga 8015  
 DB 8606 TTTAAATGATTTGTTCCAGAGCGCGGCTCTCGATATTTTAAATGATTTGTCCACGGA 8665  
 QY 8016 ggcgggctcctcgatatttaaaatgattgtgttgcagacagagcctgtgtgaaac 8075  
 DB 8666 GCGCGGCTCTCGATATTTTAAATGATTTGTTGTGACGACAGAGCTTGTGTGAAAC 8725  
 QY 8076 ccataaagcgtgccgattccgacactgcgagcgcgagatctctacacccctgcygtgta 8135  
 DB 8726 CCATTAAGACCTGTCCGATTCGCACTCGGGGCGGCACTCTCTACCTCCCTCGTGTGTA 8785  
 QY 8136 cgactgtgagccacagcgcgtgtgaataaaatccttctgtgtgtcatcaaa 8191  
 DB 8786 CCAGCTGTGGGCCACGCGCGCTTGAATTAATCTCTTGTGTGTGATCAGGA 8841

## RESULT 6

PERRY17013 7808 bp DNA VRL 14-APR-2000  
 LOCUS Porcine endogenous retrovirus gag, pol, env genes.

DEFINITION Y17013.1 GI:3005985  
 VERSION Y17013.1 GI:3005985  
 KEYWORDS env gene; gag gene; pol gene.  
 SOURCE porcine endogenous retrovirus.

ORGANISM Viruses: Retroid viruses: Retroviridae: Mammalian type C  
 retroviruses: 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 7808)  
 Czudenza, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R. R.  
 Establishment and characterization of molecular clones of porcine  
 endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)

JOURNAL MEDLINE 20219394  
 2 (bases 1 to 7808)

REFERENCE Toenjes, R. R.  
 Direct Submission

JOURNAL Submitted (30-MAR-1998) R. R. Toenjes, Paul-Ehrlich Institut,  
 Paul-Ehrlich Str. 51-59, P.O. Box 1740, D-63207 Langen, FRG

FEATURES Location/Qualifiers  
 1..7808 /organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"  
/tissue\_type="kidney"  
/cell\_line="PK15"  
260..1834  
/gene="gag"  
260..1834  
/gene="gag"  
/codon\_start=1  
/protein\_id="CAA76581.1"  
/db\_xref="GI:3005986"  
/db\_xref="SPTREMBL:O73504"  
/translation="MGQVTPPLSLTLDHMTVEBSRAHNLISVQKKGPMWQFCASEMP  
TFDVGWPESEGFENSEILAAKAIIFQTPSSHPQDEYIITWODLADPPWVPMN  
KPKRGPRILALGKKNHSAKVEPSPRIYPIEIEPTWPEPQVPPVPAQAVG  
PSAPGAVVGPAGTSPRRGATPERTDEIALLPRTYGPMPGGOLOPLQVPESS  
ADLYNMKNHPFSEDPQRLTGIVSEISMQNPWDCOOLLOTLPABERERITLAA  
RKNVPGADGRTQLONETDMGFPLTRGMDYINAGRESLKIYQALVAGLRGRSRP  
TNLAKREVMQWNPPEPSPLELMEAFRRTPDPPTSEAKASVALAFGOSLDIR  
KILQRLQELQAEILDLVREAEKAYGRETDEEKEQKEREERERERDRROBKNT  
KILAAVVGKSSRRERDFRIRSGPQSGNLGNRPDLKDQCAVCEKKGHMANRNC  
KGNKPKVLALEDKD"

2120..4726  
/gene="pol"  
2120..4726  
/gene="pol"  
/codon\_start=1  
/protein\_id="CAA76582.1"  
/db\_xref="GI:3005987"  
/db\_xref="SPTREMBL:O73505"  
/translation="MGAQISFEGRPPEVSVNKNRPITVLTLQLDDEXRLYSPPQKPPD  
IOSWLEFPQAMETAGMLAKQVPOVQLKASAPVSROYPLSREAREGIMPHVO  
RLIOGILVPOSPWNTPLVPRKPGNDYRPODLREVNKRYODLHPVPMNVLIS  
ALPERMNTVTLVDKDAFCLRLHPTSOPLFAEMRDPGRTGQTLNTRLPQFNKS  
PTIDEALHRLDANERIOHPVLTLOYVDLLAGATKODCLEGTALLLEISDLKYR  
ASAKKOICREVTYIISLKGQRMWLETKKTVQIIPAPPAKQVREFLGTAGPCR  
LWIDFATLAPLYPLTKEGEFEMAEHOKADAIKALLSAPALADPVTPTFLY  
VDERKGVARGLVLTQIGPMRRPVAVLSKIDPVSAGPVLKAIAAVALIKRADLT  
LGONITVIAHVALENTIROPDRMNTAKMTIHOSILPERVAPALNAPATILPE  
ETDEPVTCHDCHOLLIEETGVKRLDTIDPIGEVLTPTDSSIVYGGKRMAGAAVVDG  
THITMASIPBEGSAOKAEMALTQALRLAEKSNITVDSRYAFATVHGAIVYQR  
GLTSAERETKNEELISLEALHLPRKLAIHCPGHAKADLSRNGMADVAKOA  
AOAVNLPIETPRAPRROYTLEMOEIKIDQSETEGTCYSYKELIPHEG  
LEYVOOHLRLHGTKHLQOLVTSYVHLRLEGVADSVYKHCVPOLVANSRIRP  
GKRLRSGHGAHWEVDETEKPAKYGNKYLIVVDFFSGWEAVPIPKKETSYVAKY  
WRKFRGSHLEYLR"

5294..7267  
/gene="env"  
5294..7267  
/gene="env"  
/codon\_start=1  
/protein\_id="CAA76583.1"  
/db\_xref="GI:3005988"  
/db\_xref="SPTREMBL:O73506"  
/translation="MHPILSMRLPTRGGEPRRLRIPLSFIAMFLTITTPQASSK  
RLIDSNPHRPLSLIWLIDPDTGVNSTRGVAPGTWPELHCLRLINAVASTP  
PNLVRSYGFCPEGTEKEKYGCGSESEFCRRKSCVTSNDGDKMKLISODRPFSEVN  
SGGKYKVMKLYKDKSCSPSDLYLKISPTREKSGKENTOKMKNMSKGVFRTKGGGA  
GSTITRLRIETGCTPRPVAVGPDKVLAEBGPALPEPHLVPQLTSLRPTDTPSN  
GTTGLIPTNPKNSPGVPKVTQRLSLIIGAFOAINSDIPATSSCMCLSGSPY  
EGMAKGRGNVTRKHNOCOTGSRNKLITVEYSGKTCIGKAPSHOHLCTSTVYEO  
ASENDYLVGYNNWACNTGLTFCVSSVFNOSKPCVAVQIVPVYVHPEVVLIDEX  
DYRNNRPREPVSILAVMLGTAVGAGTGTAAALITGPQOLEKLTGELHAMMTDRL  
ALEESYLNLESLTISLVYIQRNGIDILPLRREGUCALAEBCCFYDHSIGATIDS  
MSKIRERLRRRRERADOGWFEGMFNPSPMWTTLLSALTGLVLLLVGPCLLIN  
RVAATREVVAGVQIMVLRKQYQGLISGEMDL"

BASE COUNT 2077 a 1971 c 2047 g 1713 t  
ORIGIN

Query Match 93.1%; Score 7644; DB 14; Length 7808;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 7746; Conservative 0; Mismatches 55; Indels 9; Gaps 6;

QY 329 atcagcagacgtgcttagagagatcatcagcgtccacccttgaggagacgcccggaggtg 388  
Db 1 ATCCACACAGCTGCTAGAGAGATCACAGGCTCCACCCTGGGGAGCCCGGGAGGTG 60  
QY 389 ggaagcgaagcagcagcctggtgtctcctactgtcgtcagtagagacagcttctgtgt 448  
Db 61 GGAGAGCCAGGAGGAGCCTGGTGTCTCTTACTGTGCTCAGAGAGACCGAGCTTCTGTT 120  
QY 449 gaaagcgaagcttcccccctccgagccgctccagctcttctgctgttgaggagacg 508  
Db 121 GAAGCGAAGAGCTTCCCTCCGCGCGCTCCGAGACTCTTTTGCCGTGTGGAAAGACCG 180  
QY 509 gaagcgtcgtgtgtgtgtgtatcgtgtgttctgtctgtgtgtgtgtgtgtgtgt 568  
Db 181 GACGGGT 240  
QY 569 cctgtctacagtttaataatgagacagagagagacaccccttaagttgtgtctgtga 628  
Db 241 CCTGTCTACAGTTTAAATATGAGACAGACAGTACAGACCCCTTAAGTTAACTCTCA 300  
QY 629 ccattgtcagtgagtgatagatccagaggtcatalaattgtcagtgatgaagtagacc 688  
Db 301 CCATTGAGCTGAGATTAGATTCAGGCGCTCATATTTGTGAGTTCAGGTTAAGAGGAG 360  
QY 689 ttggcagacttctgtgtccctgtaatgagcacaatcagatgtgtgagtgccatcaga 748  
Db 361 TTGGCAGACTTCTGTGCTCTGCTCAATGAGCCACATTCGATGTGTGAGTGCATCAGAG 420  
QY 749 gaccttaactcgaatcatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 808  
Db 421 GACCTTAAATTCGAATTAATTCCTGCTGTAAAGGCAATATTTCAACTGAGACCCAG 480  
QY 809 ctctcatcctgatalcagagacccatatactctacgtgtgcaagatttgagcaagatctcc 868  
Db 481 CTCTCATCTCGATCAGAGAGCCCATATCTTACGTGCAAGACTTGCGAAGATCTCTC 540  
QY 869 gccatggttaaacatcgtgtcaataaacaagaagccaggttcccgatctgtgtct 928  
Db 541 GCATGTGGTTAAACCATGTGCTTAATTAACCAAGAAAGCCAGGTCCCGAATCTGTGCT 600  
QY 929 ttgagagaaataaacaactcgtgagcaaaagttagagcccttctcgtatctaccg 988  
Db 601 TTGGGGAAGAAACAAACACTGCGCGCAAAAGTCAAGCCCTC-TCCTGTACTCTACCCG 659  
QY 989 agatcgaagagccgagcacttggccggaaccccaactgttcccccaccccttcaag 1048  
Db 660 AGATCGAGAGAGCGCGGAGATTGGCCGGAAGCCCAACCTTTCACACCCCTTATTCAG 719  
QY 1049 cacaggtgtgtgtgagggga-cctctgcccctcctgtgagctcgtgtgtgagggactg 1107  
Db 720 CACAGGCTCTGTGAGGGAGCCCTCTGCCCTCTGAGACTCGGTGTGAGGAGACTG 779  
QY 1108 ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1167  
Db 780 CTGCGCGAGCTGAGAGCCGAGAGGCGGACCCCGAGCGGACACACAGATTCGGGAT 839  
QY 1168 tacgcgtgcgacctaataatgagcagcagcagcagcagcagcagcagcagcagcagc 1227  
Db 840 TACCGGTGGGACCATATGCGCTCCCATGCGGGGGGCGCAATGAGACCCCTCCAGAT 899  
QY 1228 ggccttctctgtcagatctctataatgtgaaactaacatcccttctcgtgag 1287  
Db 900 GGCCTTTTCTCTCAGATCTCTATATGTGAATAACATCAATCCCTTTCTCGAGG 959  
QY 1288 atcccaacagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1347  
Db 960 ATCCCAACAGCCTCACGGGGTGTGTGAGTCTTATGTCTTACACAGCCTACTGGG 1019  
QY 1348 atgatttcacaagctgtgcagacaactctcacacaagcagcagcagcagcagcagcagc 1407  
Db 1020 ATGATTTCACACAGCTGTGAGACACTCTTTCACAGCCGAGAGCAGAGAAATTCGT 1079  
QY 1408 tagagctagaaaaaagtctcctgtggtccgagcagcagcagcagcagcagcagcagcagc 1467



1080 TAGAGGCTAGAAAAATCTCTGGGCGGACGGGCGACCCACGACTTGCATAATGAGA 1139  
1468 ttgacatggattcccttgatcgcccggttggaactacaacagcgctgaagtagg 1527  
1140 TTGACATGGATTTCCTTCCCTTGACTCGCTCGGCTTGGAATCAACAGCGCTGAAGGTAGG 1199  
1528 agagctgaaatctatcgccaggctctgtgtgcggtctccgggagcgctcaagagcg 1587  
1200 AGAGCTGAAATCTATGCGCAGGCTCTGGGCGGCTCTCCGGGGCGCTCAAGAGGCG 1259  
1588 ccactaatcttgctaaagtaagaaagtgatgacggagccgaagaaactccctcgat 1647  
1260 CCACATAATTGGCTAAGGTTAAGAGGATGATGACAGGACCAAGCAACCTCCCTCGGTAT 1319  
1648 ttcttgaaagcctcatgaaagccttcaggcggttcaccccttttgatcctcaactcagag 1707  
1320 TTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGTTACCCCTTTTGATCTCTACCTCGAGG 1379  
1708 cccagaagcctcagtgcccttgctcattgggcagtcggtctctgatacaggaag 1767  
1380 CCCAGAAAGCCTCACTGGCCCTGGGCTTCATTTGGGCACTCGGCTCTGGATTCAGAAAGA 1439  
1768 aactcagaagctggaaggtttacagagagcttgatcgtgactagtgagagagcag 1827  
1440 AACTTCAGAGACTGGAAGGTTTACAGAGGCTGAGTTACGTGATCTAGTGAGAGGAG 1499  
1828 agaaggtglatatcagaagggagacagaagaagaaagaaagaaagaaagaaagaa 1887  
1500 AGAAGGCGTATACGGAAGGAGAGACAGAAAGAGAGAAAGAAAGAAAGAAAGAGAGA 1559  
1888 gagaagaagaagaaagaaagctgtagagagcagaagaagaagaatctgaactctg 1947  
1560 GAGAAGAAAGGAGAGAAAGCTGATAGACGGCAGAGAGAGAAATTTGCTTAAGATCTTGG 1619  
1948 ccgagctggttgaaggaagaaagcagaaggaagaaagaaagaaatctgaagaatga 2007  
1620 CCGCAGTGGTTAAGGAAAGACAGCAGAGAGAGAGAGAGATTTAGGAAATTTAGGT 1679  
2008 cagggccttagacatcaggaagaaactgggcaatagaagcccaactgcaagaagcagctg 2067  
1680 CAGGCGCTTAGACAGTCAAGGAACTGGGCAATAGGACCCCACTCACAAGGACCGAGTGG 1739  
2068 cgtatgttaagaaaaagacactggcagaagaaactggcccaagaagaaagaaagaaag 2127  
1740 CGTATTGTAAAGAAAAAGACACTGGGCAAGAACTGCCCAAGAAAGGAAACAAAGGAC 1799  
2128 cgaagctccttagctctagaagaagaataaagattagggagagacggggttcggagccctcc 2187  
1800 CGAAGTCTCTAGCTTAGAAGAAATTAAGATTAGGGGAGAGCGGGTTGCGACCCCTCC 1859  
2188 ccgaagcccaaggttaacttgaaggtggagggcgaacagctgaagttccctgtgtgatacg 2247  
1860 CTGAGCCCAAGGGTAACTTTGAAGGTGAGGGGCAACCACTTGAAGTTCCTGTTGATACCG 1919  
2248 gaacggaagattcagtgctgtctacaacccaataggaataaagaagaaataactctgg 2307  
1920 GAGCGGAGCATTCAGTGTCTGCTACAACTTAGGAAACTTAAAGAAAAAAATCTCGG 1979  
2308 tgaatgggtgacacagggcacaagcagatcatalcatalgactacccgaagaacgltgactgg 2367  
1980 TGGTGGGTGACACAGGGCAACGCGAGTATCCATGAGCTACCGAAGAACCGTTGACTTGG 2039  
2368 gaatgggaagcggtaaacccactggttctgtgataccctgaagggcccaagtaaccctctag 2427  
2040 CAATGGAGCGGTAAACCCACTCTTCTGCTCATCCCTAGAGTCCCAAGTACCCCTTTTAG 2099  
2428 gtagagacttactgacaagaatgagagctcaaatctcttcttgacaagaagaagaccagaag 2487  
2100 GTAGAGACTTACTGACCAAAATGGAGCTCAAAATTTCTTTTGAACAAGAGACCGAAG 2159  
2488 tgcctgtaatacaaacccatacactgtgttgacccccaatagatgataatcgac 2547  
|||||

2160 TGCTGTGAATAAACAAACCCATCACTGTTGACCCCTCCAAATTAAGATGATATGAC 2219  
2548 tatattctcccaagaagagcctgatacaatatacaagctccggttgagagagttcccc 2607  
2220 TATATTCTCCCAAGTAAGACCTGATCAAGATATACAGTCTGTTGGAGCATTTCC 2279  
2608 aagcctgggaagaacacagagagatggtttggcaagaagatctcccaagatctatc 2667  
2280 AACCTGGGAGAGAACCGCAGGATGGTTTGGCAAGCAAGTCCCAAGATATATTC 2339  
2668 aactgaagcagtgctacacagatcaagtcagacagtaacccttgatagagagctc 2727  
2340 AACTGAAGGCGCAGTGTACACCGATATCAGTCAAGACAGTACCCCTTGAGTAGAGAGCTTC 2399  
2728 gagaaggaatttggcagatgttcaagaagttatccaagaagagcatcctagtctcgtcc 2787  
2400 GAGAAGAAATTTGGCCGATGTTCAAGATTAAATCCAAAGGAGATCTAGTTCCGTGTC 2459  
2788 aatcccttgaaatacctcccttgctacacggtltagaagcctggagccaatgatalcgac 2847  
2460 AATCCCTTGGAATPACTCCCTGCTACCGGTTAGGAAGCCTGGACCAATGATTAATCGAC 2519  
2848 cagtaacagacttgaagagagatcaataaagggtgcagagacataacccaaggtccga 2907  
2520 CAGTACAGGACTTGAGAGAGTCAATPAAAGGTTGACAGACATACACCCACCGTCCGA 2579  
2908 accctttaaactcttgagagccctccgctcgaagaaactgtgatacagatltgact 2967  
2580 ACCCTTAACTCTTATAGGCGCTCCCGCTACAGGAACTGTACAGATTTGGAGT 2639  
2968 taaaagatgctcttcttgcttgatgatacaccactagccaacacttttgctctg 3027  
2640 TAAAGATGCTCTTCTGCTGTGATTAACCCCACTACCAACCACTTTGGCTTCG 2699  
3028 aatgaaagatccaagfagcaggaagaaacgggagcagctcagacccgagctcccaag 3087  
2700 AATGAGAGATCCAGGTACGGAGAACCGGAGACTCACTGAGACCCGACTGCCCAAG 2759  
3088 ggtccaagaaactcccgacacatctltgacgaagccttacaagagaccttggccaactca 3147  
2760 GGTTCAGAACTCCCGACCATTTTGAAGAGCCCTACACAGGAGACTGGCCAACTTCA 2819  
3148 ggtatccaacacccctcaggttgaagcctctctcagatcgtgagatgactctctggggag 3207  
2820 GGATCCAAACACCTCAAGGTGACCTCTCCAGTACAGTGAAGTACCTCTTCTGGCGGAG 2879  
3208 ccacaaacagagactgtcttaagaagtacgaaggaactcagctgagaaatltgactag 3267  
2880 CCACCAACAGAGACTGCTTAGAAGGTACGAAGGACTACTGCTGGAATTTGCTACCTAG 2939  
3268 gctacagagcctctgctaaagaagccagatltgcaagagaagagtaacatacttgggt 3327  
2940 GCTACAGAGCTCTGTGTAAGAGGCCAGATTTCAGAGAGAGAGTAACTATCTGGGTT 2999  
3328 acaagttgagggcgagcagcagatgtgtgacagggagcaagaaagaactgagtcaga 3387  
3000 ACAGTTTCCGGGGCGGAGGATGCTGACAGGAGGCAAGGAAATACTGATCTCAGA 3059  
3388 tacggccccaacacacagcaacaagaatgagagatlttgggagacagctgagatltgca 3447  
3060 TACCGGCCCAACACAGCAACAGTGAAGAGTTTGGGAGACAGCTGGATTTTGCA 3119  
3448 gacttgatcccgagggtltgagaccttagcagcccaactacccgtaaccaagaa 3507  
3120 GACTGTGATCCCGGGGTTTGAGACTTAGAGAGCCCACTCTACCCGCTTAACCAAGAA 3179  
3508 aagggagattctcctgggtctctgagcagcagaagagagatgagatcacaagaagcc 3567  
3180 AAGGGGAATTCCTCGGGCTCTGAGCACCGAAGGCAATTTGATCTATCAAAAAGGCC 3239  
3568 tgttagcgaactctctggccctcctgacgtaactaacaacacttacccttatgtg 3627  
3240 TGCTAGCGCACTCTGCGGCTCCCTGAGGTAACTAAACCCCTTACCTTATGTGG 3289  
|||||



QY 3628 atgagcgttaaggagtagcccgaggagitttaaccccaaacccctagagaccatgagagagac 3687  
|||||  
Db 3300 ATGAGCGTAAAGGAGTACCCGAGAGAGTTTAAACCAAAACCCCTAGAGCATGAGAGAGAC 3359  
QY 3688 ctgttgccctactgtcaaaagagcttgatcctctgtagcagtggtgtgcccgtatgctga 3747  
|||||  
Db 3360 CTGTTGCTTACCTGTCAAGAGAGCTTGAATCTGTAGCCAGTGTGGCCCTATGCTGA 3419  
QY 3748 aggcatalcgacgttgtagccatctgtgtcaagagcgctgtaaaattgactttgtagcaga 3807  
|||||  
Db 3420 AGGCTATGCGAGCTGTGGCCATCTAGTCAAGAGACGCTGCAAAATTGACCTTGGGACAGA 3479  
QY 3808 atataactgtataagcccccaatgcatgtgagaaacatggttccgagccccagaccgat 3867  
|||||  
Db 3480 ATATAACGTAAATAGCCCCCATGCTCATGTGAGAACATGCTGTGGCAGCCCCAGACCGAT 3539  
QY 3868 ggaatgaacaaagcccgatagccacatataaagccctgtcttcaagagaaggtcaact 3927  
|||||  
Db 3540 GGATGACCAACGCCCGCATGAGCCCATATCAAAAGCCGTCTTCAAGAGAGGGTCACTT 3599  
QY 3928 tgcgtccaccagccgctctcaaccctgcacctcttcgcttgaaagagagatgtaaccag 3987  
|||||  
Db 3600 TCGCTCCACCAAGCGCGCTCTCAACCTTCCACTTCTGCTGAAGAGACTGATGAAACAG 3659  
QY 3988 tgaactatgattgccaatcaatltgatltgaagagactgggtgtccgaagagacttaag 4047  
|||||  
Db 3660 TGACTCATGATTTGCCATCAACTATTTGATGAGAGAGACTGGGGTCCGCAAGACCTTACAG 3719  
QY 4048 acataacgcctgaactggaagagtgtaactggttcaacttagcgaagagaagatgtagtgg 4107  
|||||  
Db 3720 ACATACCGCTGACTGAGAAAGTGTAACTGCTTCACTGAGAGAGAGAGATGATGTGGTGG 3779  
QY 4108 aaggtaaagagatggtctggggcgagtggtgtagacggagccgacagatctgggccaaga 4167  
|||||  
Db 3780 AAGGTAAAGAGATGGCTGGGGCGGGGTGTGACGGAGCCACACAGATCTGGGCCAGCA 3839  
QY 4168 ggcctggcgaagaaacttgagcgcaaaagctgaactatgagccctcaagaaagcttggc 4227  
|||||  
Db 3840 GCTTGCGGGAAGGAACTTCAAGCCCAAAAGGCTGAGCTCAAGGCCCTCACCAAGCTTTTGC 3899  
QY 4228 gactggcgaaggaatccataaacaattatagcgaagcagatgtagccttggagagc 4287  
|||||  
Db 3900 GGCTGGCGGAAGGAATTCATTAACATTATATGAGAGACAGCTATGCTTTGGCAGCTG 3959  
QY 4288 caacatgacacggggccatctataaacaaggggtgtgttactcaagcagagggagaa 4347  
|||||  
Db 3960 CACACGTACACGGGGCCATCTATAAACAAAGGGGTGTGTACTCAGCAGGAGGGAGAA 4019  
QY 4348 taagaacaagaagaattctaaagctatagaagcccttaacttggccaaaagctag 4407  
|||||  
Db 4020 TAAAGAAACAAAGGAATTTCTTAAGCCTTATTAAGCCTTATTAAGCCTTATTAAGCCTT 4079  
QY 4408 ctattataactgttctgtagacatcagaagaagccaaagatcatalctagaaggagacaga 4467  
|||||  
Db 4080 CTATTATACACTGTCTCTGAGCATCAGAAAGCCAAAGATCTCATATATTAAGGAAACAGCA 4139  
QY 4468 tggctgacccgggtgtgccaagcagggagcccaaggtgtttaaactctgactataatagaa 4527  
|||||  
Db 4140 TGGCTACCGGGTGTGGCAAGCAGGCGCCAGGCTGTATTAACCTTGTGCTTATTAATAGAA 4199  
QY 4528 cggcccaagccccagagaaccagacagtagacccttagaagaagctgagagaagataaaaa 4587  
|||||  
Db 4200 CGCCGAGAGCCCGAGAACCCAGACGACATACCTTAAAGACATGAGGCAAGGATTAAGAA 4259  
QY 4588 agatagaccagttctctgtagactccggaaggagactgtatatacctcataatggagaa 4647  
|||||  
Db 4260 AGATAGACCAAGTCTCTGAGACTCCGAGAGGAGCTGTATATCATATGAGGAAAGAA 4319  
QY 4648 tccctggcccaagaagaaggttagaataatgtcccaagatacatcgtctaaaccaactag 4707  
|||||  
Db 4320 TCTGTCCCAACAAAGAGGGTTAGATATGTCCAAACAGATACATGCTCAACCCACCTAG 4379

---

QY 4708 gaactaaacccctgcagcagttgtgtcagaacatcccccttaataatgttctgagctacag 4767  
|||||  
Db 4380 GAACATAACACCTGTGACACACTTGTGTAGAACATCCCTTATCATGTTCTGAGGCTACCG 4439  
QY 4768 gagtgctgacatcgtgtgtcaaaacattgtgtgccttcagcagctgtgttaatgtaactct 4827  
|||||  
Db 4440 GAGTGGCTGACGTGGTGTCAAAACATGTGTGTGCTGCGCAGCTGTTATGTGAATCTCT 4499  
QY 4828 ccagaataactccaggaagaagactaaggaggagcccaacccagcgctcaactgtggaatg 4887  
|||||  
Db 4500 CCGAATACCTCCAGGAAGAGACTAAGGGAGGCCACCCAGCGCTCACTGGGAATGTG 4559  
QY 4888 acttcaactgagtaaaagccggtctaaatagcgaacaaatatactatgtgttctgtagaa 4947  
|||||  
Db 4560 ACTTCACTAGATGAAGCCCGCTAAATATGGAACCAATATCTATTGTTTGTAGACA 4619  
QY 4948 cctttcagaatggttagagagcttaccctaaagaagaagacttaacacgtgtgtgta 5007  
|||||  
Db 4620 CCTTTCAAGATGGGTAGAGAGCTTATCTTAAGAAAGAGACTCAACCGTGGTGGCTA 4679  
QY 5008 agaaaaactgtagaggaatttttccaagatttggaatcactaagtaataaggtcaaga 5067  
|||||  
Db 4680 A-AAATATCTGAGGAAATTTTCCGAGATTTGGAATACCTAAGGTAAATCCGGTACAGA 4738  
QY 5068 atgltcagacttgcgttgcagtagaagtcagggactgccaagatatttgaggtatg 5127  
|||||  
Db 4739 ATGCTCACACTTTTGTGTGCGCAGATAGCTAGAGGACTGGCCAAAGTATTGGGATGTAT 4798  
QY 5128 ggaactgcatgtgtcatcagaccccaagctcagacaggttagagagatgaa 5187  
|||||  
Db 4799 GGAATCTGATTTGTGATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGATGA 4858  
QY 5188 ccattaaagaagcccttaccacaaatagacacagaagacgtggaatattatgtagtgc 5247  
|||||  
Db 4859 CCATTAAGAGACCTTATACCAATTTGACACAGACCTGCAATTAATATTGATGATGATG 4918  
QY 5248 tccgtcccttgtgtctttaggttaggaacacccctcagacagtttgggtgcacccct 5307  
|||||  
Db 4919 TCTGCTCCCTTTGTGCTTTTATAGGTTAGGAACACCCCTGGAAGTGTGCTGATGACCCCT 4978  
QY 5308 atgaattgctctcaagggggagaccccccgtgtgcaagaattgtccttgcacatgtgctg 5367  
|||||  
Db 4979 ATGAATGTGCTTACGGGGAGCCCCCGCTTGTGAAGAAATGCTTCTATCATATGTGCTG 5038  
QY 5368 atgtgctccttccagaccttgtctcttagcctaaagggcctgaagtggttagaggaagc 5427  
|||||  
Db 5039 ATGTCTGCTTTCCAGGCTCTGTCTTCTAGGCTCAAGGCTCTGAGTGGGTGAGCAAC 5098  
QY 5428 gaggctggaagcagctccgggaagcctactcaag--aggaagactgcaagtlccacatc 5484  
|||||  
Db 5099 GAGCTGTGAACAGCTCCGGAGGCGCTACTCAGGAGAGAGACTTGCATTCACATC 5158  
QY 5485 gcttccaaagttaggaatcagctatgtttagacgcaaccggtgcaagaaactcgaaactc 5544  
|||||  
Db 5159 GCTTCAAGTGTGGAATTCATGTATGTATTAAGCCCAACCGTGCAGGAACCTTCAGACTC 5218  
QY 5545 ggttgaaggaacctatctgcttactttagacacaacaagcgctgtgaaagtgcgaagaa 5604  
|||||  
Db 5219 GGTGAAGGGCCCTTATCTGTACTTTGACCAACACCAACGCGTGTAAAGTGCAGAGAA 5278  
QY 5605 tcccaactgtatccatgcatcccaagltlaagcggcgccaactccagactcgggtgtga 5664  
|||||  
Db 5279 TCTCCACCTGTGATCATCATCCACGTTAAAGCTGGGCGCACCTCCGACTCGGGGTGGA 5338  
QY 5665 gagccgaagaagactgagaatcccttaagcttgccctccatgcttgcttcttactcta 5724  
|||||  
Db 5339 GAGCGGAAGAGACTGAGATTCCTTTAAGCTTGCCTTCATCGCTGTGTTCTTACTACTA 5398  
QY 5725 acaataactccccagcgccagtagtaaacgcttataagacagcttgcgaacccccataagcct 5784  
|||||  
Db 5399 ACAATTAACCTCCCGAGGCGAGTAGTAAGACGCTTATTAACAGGCTGGAACCCCATAGACT 5458  
QY 5785 ttatcccttactgctgcatatattgacccctgatacgggtgttcaactgtlaaataagactga 5844

|||||  
Db 5459 TTATCCCTTACTGGCTGATTTATTTGACCTGATACGGGTGCTGACTGTAATATACACTCGA 5518  
QY 5845 ggtgtgtcctcctagaagcaccctggtgctggaactgacatcttgcctccgactgathaac 5904  
Db 5519 GGTGTGCTCCTTAGAGGACACTGGTGGCTGAACCTGATTTCTCCCTGATGATTAAC 5578  
QY 5905 cccgctgttaaaagcacacacctcccaactagtcgtagttagtgggtctctgtctgcca 5964  
Db 5579 CCCGCTGTTAAAGCACACCTCCCAACCTAGTCCGTAAGTATGAGGTTCTATTCCTGCGCA 5638  
QY 5965 ggcacagaaagaaatactatggtgggtgtctggaatccctctgtagagaatgagc 6024  
Db 5639 GGCAACAGAAAGAAATACTGTGGGGTTCTAGAGAACTCTTCTGTAGAGATGAGAC 5698  
QY 6025 tgcgtcacctcccaagcttggaacttggaatgycgaatctctctccagacccggtlaaa 6084  
Db 5699 TGGCTGACCTCCACGATGAGACTGGAAATGGCCGATCTCTCCAGGACCGGATAAA 5758  
QY 6085 ttctccttgcattcctcggcgggaagtaacaaatgtagaactataaagataag 6144  
Db 5759 TTCTCCTTTGTAATTCGGCCGGGCAAGTACAAAGTGAATATATAAGATAG 5818  
QY 6145 agctgtcccatcagaactagatlatcctaaagaataagttcactg--aaagaaacag 6202  
Db 5819 AGTGCTCCCATCAGACTTGAATTAATTAAGATTAAGTTCACTGAAAAAGAAACAG 5878  
QY 6203 gaaatattcaaaagtgataataatgtagatgcttgggaatagttttat-tatatgac 6261  
Db 5879 GAAATATATTAAGATGATTAATGATGATGAGTGGGAAATGTTTTTATTAATTAAGGC 5938  
QY 6262 ggggaagagaggtcacttaacatccgcttagagaagaagggagacacccct 6321  
Db 5939 GGGGAGAGAGGTCCACTTTAACATTCGCTTAGGATAGAGACGGGACAGAACCCCT 5998  
QY 6322 gtggcaatgggaaccgataaagtaactgctgtaacagggcccgccctggaacacg 6381  
Db 5999 GTGGAGTGGGACCCGATTAAGTACTGGCTGAACAGGGGCCCGGCGCTGAGACACCG 6058  
QY 6382 cataactgcccgtgcccnaatlaactcgtcgtcggcctgacataaacaagccgctagc 6441  
Db 6059 CATAACTTGGCCGTCGCCCAATTAACCTGCTGGCGCTGACATTAACACACCCGCCCTAGC 6118  
QY 6442 aacgtaccactgtagatctcctcaacaacgctagaactcccaaggtgtccgt 6501  
Db 6119 AACGTACCTGATGATTAATTCCTACACAGCCTTAACAACTCCCAAGGTCTCTGTT 6178  
QY 6502 aagacaagacagagactctcagctcattccagggagcttccaaagcatcaactcaac 6561  
Db 6179 AAGACAGAGACAGAGCTTTCAGTCTCATCCAGGAGGCTTCCAAAGCATCACTCCACC 6238  
QY 6562 gaacctgtagcgaactctctgtgtgtccttgcctacacgaagcctcctattatgag 6621  
Db 6239 GACCTGATGCTCCTTCTTGTGTGGCTTGTCTATCTCCAGGCGCTCTTATTAATAG 6298  
QY 6622 gggatggttaagaaagaaatlaaagtgtgcaaaagacatagaatcaatgtagatg 6681  
Db 6299 GGGATGGCTTAAGAGAGAAATTCATGTGACCAAGAGCATTAAGAAATCATGTACATGG 6358  
QY 6682 ggggtccgaaataagctcactcaactgaagttccgggaaaggggacatgcatagaana 6741  
Db 6359 GGGTCCGAAATTAAGCTTACCCTCAGTGAAGTTCCGGGAGGGGACATGATGAGAAA 6418  
QY 6742 gctcccccatacccaacaacacttgcctatagtagtctgtgttatgagcagcctcga 6801  
Db 6419 GCTCCCATCCACCAACACCTTGTATAGTACTGTGTATGACAGAGCCCTCAGAA 6478  
QY 6802 aatcagattatagtagcctggtatataacaggtggtgagcatalactaggtttaaacc 6861  
Db 6479 AATCAGTATTTAGTACCTGGTTATACAGGTGGTGCATGCAATACCTGGTTAACCC 6538  
QY 6862 tgtgttccacctcagcttcaacacatccaaagatttctgtgcatgtgtccaaatgct 6921  
|||||

Db 6539 TGTGTTCCCTCCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGATGCTCAAAATCTGC 6598  
QY 6922 ccccgagtgtagtaccatccctgaagaagtgtgctttagatgaaatagtagtaccatgataac 6981  
Db 6599 CCCGAGTGTACTACCATCTGAGAGAGTGGTCTTGATGAAATATGACATATGATTAAC 6658  
QY 6982 cgaacaaagaaagacccglatcccttaccctagctgtlaatgctcgtgaatagaagcgc 7041  
Db 6659 CGACCAAAAAGAAAGAACCCGATATCCCTTACCTAGCTGATATGCTGAGATAGGAGCGGC 6718  
QY 7042 gttgctgtagaagcaggaagcagctgacctgatacaagagacacagcagtagaaga 7101  
Db 6719 GTTGGCGTAGAAACAGAGGACAGCTGCCGTGATCACAGACCCACAGAGCTTAAAGAAAGGA 6778  
QY 7102 ctgtgtgactacatgctgacatgaaagaagttccggccttaagaaggtctgtagc 7161  
Db 6779 CTGTGGACTATATCGGCTGATGACAGAGATCTCGAGCTTAAGAGAGTCTGTTAC 6838  
QY 7162 aacctagaagatccctgactctctgtctgaagtggtctcaagaacccggagggatga 7221  
Db 6839 AACCTAGAAAGAGTCCCTGACTTCTGTGAAGTGGTTCTACAGAACCGAGGGGATTA 6898  
QY 7222 gactgtcttctcagaagaaggtggtatgtgcaagccttaagaagaatggtgttc 7281  
Db 6899 GATCTCTGTTCTTAAGAGAGAGTGGTTATGTGACGCTTAAACAAAGAAATGTCCTTC 6958  
QY 7282 tatgtatgactcagagagccatcagagactccataaacaagcttagaanaaagttagag 7341  
Db 6959 TATGTGATCCTCAGAGAGCCATCAGAGACTCCATAGCAAGTTTGAAGAAAGTTAGAG 7018  
QY 7342 aggcctcgaaggaagagagcgtgacacaggggtgttgaagaatggttcaacaggtct 7401  
Db 7019 AGCGTCGAAGGGGAAAGAGAGGCTGACAGGGGTGTTGAAGAGATGTTCAACAGTCT 7078  
QY 7402 ccttgatgacacacccctgtcttctgtcgtgacggggccctagtagtctgtcccgta 7461  
Db 7079 CTTTGATGACCAACCTGCTTCTGTGACGGGACCCCTATATCTGCTGCTGTTA 7138  
QY 7462 cttaagttggcctgtcttaataatgagttgtgtgcttgttagaagaacagtgagt 7521  
Db 7139 CTTACAGTTGGCTTGTCTTAATTAATAGTTGTGCTTGTGTAGACAGAGTGGGT 7198  
QY 7522 gcactccagatcactgttagtagcaacagtagtaacaaagccttctgacaaagagaact 7581  
Db 7199 GCAGTCCAGATCATGTGACTTAGGCAAGATACCAAGGCTTGTGACCAAGAGAACT 7258  
QY 7582 gaacctagcctcccaagctcctaagattgaactatlaacaagaacaaaggtggygaag 7641  
Db 7259 GACCTTAGGCTTCCAGATTAAGATTAAGATTAATTAACCAAGCAAGAGTGGGAATG 7318  
QY 7642 aaagatgaataatgcaacctaaacctcccaagacccaggaagttaataaagctctaa 7701  
Db 7319 AAAGATGAATATGCAACCTTACCTTCCGAAGCCCAAGAAAGTAAATAAAAAGCTTAA 7378  
QY 7702 tgcccccgaattacaagacctgtgctgctgcaagtaaatgtagaagttcacactccta 7761  
Db 7379 TGCCCCGAATTCACAGACCTGCTGCTGCTGCAATTAATAGTGAAGAGTACACTTCTTA 7438  
QY 7762 ttgttccagggcctgtctatccgtgctgaagtagataaagaagaagtgtgctatag 7821  
Db 7439 TTGTTCCAGGCGCTGCTATCTTCTGCTTAATTAAGATTAAGCAAGAAATGAGTGTATG 7498  
QY 7822 ctatctggaattctgaataactgacagcacaatagaagaattgattacatctgacgc 7881  
Db 7499 CTTATCTGATTTCTGTAAACTGACGTGACCAATTAAGAAATGATTAACATTTGACAG 7558  
QY 7882 cctagtagacctatcaactgcaactgttcaactgtcccaagagcccaagcagatgag 7941  
Db 7559 CTTAGTACCTATCTCAATGCAATCTGTCACTCTGCCAGGACCCAGCATATCGGA 7618  
QY 7942 cctccgagatattttaaaatgtagtgcacaggaagcggggctcgtatattttaa 8001  
Db 7619 CTTCCGAGCTATTTTAAATGATGTGCTCCAGGAGCGGCTCTGATATTTTAAAT 7678  
|||||



301 ctcacacgtaagagctgagagactgtatcacagcagctgtaggaagatcacagctg 360  
|||||  
828 CTCACACCGTAAGAGCTGAGAGACTGTATCAGCAGAGCTCTAGAGAGATCACAGGCTG 887  
Db  
361 ccaacctggggagacgccccgggaggtggggagagccagagcctgtgtgtctctact 420  
|||||  
888 CCACCCCTGGGGAGCGCCCCGGGAGGTGGGAGAGCCAGGAGCGCTGTGTCCTACT 947  
Db  
421 gtctgttagagagacgagttctgtgtgtgaagcgaagctccccctccggccggtccg 480  
948 GTCTGTAGAGAGACCGAGTTCGTGTGTGAAGCGAAAGCTTCCCTCCGGCGCTCCG 1007  
Db  
481 acctcttgccgtctgtggagagacgagcgggtcgtgtgtctgtatctgtgttt 540  
1008 ACTCTTTTGCTGCTGTGTGGAGAGCGGAGCGGCTGTGTCTGTATCTGTGGTTT 1067  
Db  
541 ctgtctgtgtgtctgtctgtgtgtgtgtctgtctgtctgtctgtctgtctgtctgt 600  
1068 CTGTTCGAGTGTCTGT 1127  
Db  
601 tgaactaccccccttaagtttagactctgaacattgagctgaagtttagatccaggtcata 660  
1128 TGACTACCCCTTATGTTGACTCTGACACATTGGACTGAAGTTAGATCCAGGGCTCAT 1187  
Db  
661 atttcaagttcaggtttaagaaaggaccttggcagacttctgtgacctgaatggccaa 720  
1188 ATTTGTCAAGTTCAAGTTAAGAAAGGACCTTGCGACACTTCTGTGCTCTTAATGGCCAA 1247  
Db  
721 catctgattgttagatgagcaltcaagagggaccttaattcgaatlaacctgtgtgt 780  
1248 CATTCGATGTTGATGAGGCAATCAGAGGAGGACCTTAAATTCGAATTAATCTGTGCTGTT 1307  
Db  
781 aggcacatcatttttcaagcttgagaccgggtctcatcctgtatcaggaagccatataccta 840  
1308 AGGCATCATTTTTCAGACTGAGACCGGCTCATCTGATCAGAGACCTTATATCTTA 1367  
Db  
841 cgtgcaagatttgacaaagaactcctccgcatggtgttaaacatgtgtaataataacaa 900  
1368 CGTGCAAGATTGACAAAGATCTCTCCGCAATGGGTTAAACCATGCTTAATTAACCAA 1427  
Db  
901 gaaagcagagtcctccgaatcctgtgtctgtgagagaaanaaacacacgcgcgcaaaag 960  
1428 GAAAGCAGAGTCCCGAATCCTGCTGTTGGAGAGAAAAACAACACTCGGCCCAAAAG 1487  
Db  
961 tctgagccctctctcgtatctatcacccagagatcgagagacgcgcgacttggccggaaccc 1020  
1488 TCGAGCCCTTC-TCTCTGATATCTACCCGAGATCGAGAGCCGCGACTTGGCCGGAACCC 1546  
Db  
1021 caactgtttcccccccccttatcacagacaggtgtgtgtgagggga-cctctgccccct 1079  
1547 CAACCTGTTCCTCCCAACCCCTTATCTCAGACACAGGGTGTGTGAGGGGACCTCTGCCCCCT 1606  
Db  
1080 cctgagactccggtgtgtgagagacgtgtcgtccggagactcgagccggaagagcgccacc 1139  
1607 CCTGAGAGCTCCGAGTGTGGAGAGACTGCTGCGGAGACTCGAGGCGGAGAGGCCCAAC 1666  
Db  
1140 ccgagagcgagacagacagatcgagatlatcacgctgtcgacactatgacctccatgca 1199  
1667 CCGGAGCGGAGACGAGATCGCATATTACCGCTGCGCACCTATGCGCCCTCCCATGCCA 1726  
Db  
1200 gggggcgaatttgcaagccccctcagatattggcccccttctctctgagatctctataatgg 1259  
1727 GGGGGCCAAATTGCAACCCCTCCAGTATTGGCCCTTTCTTCTGCGAGATCTATATAATGG 1786  
Db  
1260 aaaaataacacatcccccttctcgagagatccccagacgctcacaaggttggtagtcc 1319  
1787 AAAACTAACATCCCTCTTCTGAGAGATCCCAACGCTCACGGGTTGGTGAAGTCC 1846  
Db  
1320 ctatgtctctcacagactacttggatgattgttcaacagctgtctgcagacactctc 1379  
1847 CTTATGTTCTCTCACAGACTACTTGGATGATGTCAACAGCTGCTCAGACACTCTTC 1906

1380 acacccgagagcgagagagaattctgttagagcttagaanaaaatgtctctggcgccac 1439  
1907 ACAACCCAGGAGGAGAGAAATTCTGTAGAGCTTAGAAAAAATGTTCTCGGGCCGAC 1966  
Db  
1440 gggcgacccagcagatgtgcaaaatgagatlgaaatggatcttccctgtactgccccgt 1499  
1967 GGGCGACCCAGGAGTTCGAAATGAGATGAGATGGATTTCTTGTACTCCCGCGGT 2026  
Db  
1500 ttggactacacacgctgtgaagtgaggaagcttgaatctatccagagctctgtgt 1559  
2027 TGGAGCTACAAACCGCTGAGGTAGGAGAGCTTGAATAATCTATGCCAGGCTGTG 2086  
Db  
1560 gcggtctccggggcgctcacaagcgcccaactaatttggctlaaggtlaagagagtgatg 1619  
2087 GCGGCTCTCCGGGGCGCTCAAGCGGCCCACTAATTGTGCTAAGGAAGAGAGTGATG 2146  
Db  
1620 caagagacgagacgaacctccctcgtgtattctctgagagctctatlgaaagcttcagcgg 1679  
2147 CAGGAGCCGAAACGAACCTCCCTCGATATTCTTGAGAGGCTCATGAAAGCTTCAAGCG 2206  
Db  
1680 ttcaaccttttgatctctactcagaagggccagaagaagcctcagttggccctgtgctcat 1739  
2207 TTCACCCCTTTTGATCTTACTCTCAGAGGCCCAAGAGCTCAATGCTGCTTCATT 2266  
Db  
1740 gggcagtcggtctctgatalcagaaagaaacttcagagactggaagaggttatacagagact 1799  
2267 GGGCAGTCGGCTCTGATATCAGAAAGAACTTCAGAGACTGGAAGGTTACAGAGGCT 2326  
Db  
1800 gaattcagtgatcagtgagagagcgagagaaggtgaltacagaagggagacagaag 1859  
2327 GAGTTACGTATCTAGTGAAGAGAGCGAGAGAAAGTGTATACAGAAAGGAGACAGAAAG 2386  
Db  
1860 gaggagaaacagagaaagaaagagagagagaaagagagaaagacgtlgaataacg 1919  
2387 GAGAAAGAAAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2446  
Db  
1920 caagagagaagatttgactaagatctgtgcgcagtggttgaagagagagacagagag 1979  
2447 CAAAGAGAAATTTGATTAAGATCTTGCCGAGTGGTTGAAGAGAAAGACAGAGAGAG 2506  
Db  
1980 agagagagagatttgaagaaatlaagtcagcccttagacagtcagggaaacttggcaat 2039  
2507 AGAGAGAGAGATTTAGAAATTAAGTGTAGCCCTTAGACAGTCAAGAGAACTGGGCAAT 2566  
Db  
2040 aggaacccacttgacaaagagccagtggtgtatgtgaagaaagagacatgtggaaag 2099  
2567 AGGACCCCACTCGAAGAGGACAGTGTGCTATTGTAAAGAAAGAGACCTGGCAAG 2626  
Db  
2100 aactgcccagaagagagaaacaaagagacgaagctcctagctctagaagagaataaagat 2159  
2627 AACTGCCCAAGAGAGAAACAAAGAGCCAGAGTCTTACTTAGAAGAGATTAAGAT 2686  
Db  
2160 taaggagagaggttctgagacccccctcccgagcccaaggtlaacttgaagtgtgaggg 2219  
2687 TAGGGAGAGCGGGGTTGAGACCCCTCCGAGCCCAAGGTAATCTTGAAGGTGGAGGG 2746  
Db  
2220 caacacgttgatgtcctgtgtgtatcacggagcgagagattcagatgtctgtctacaacatla 2279  
2747 CAACCACTTAAAGTCTGTTGATACCGGAGGAGGAGATTCACTGCTGTACCAACATTA 2806  
Db  
2280 ggaagaaactaaagaaaaaatacctgtgtgtatgtgtccacagagcaagagcagatca 2339  
2807 GGAAGAACTTAAAGAAAAAATCTGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2866  
Db  
2340 tggactacccgagaaacgtttagcttggagtgagagcggglaaoccatcgtttctgtgtc 2399  
2867 TGACTACCCGAAAGAACCTTGAATGTGAGTGGAGAGGAGTAACCCACTGTTCTGTGTC 2926  
Db  
2400 atccctgagtgccagagtaacccctctctgtgtgagagacttactgtgccaagatggagatc 2459  
2927 ATCCCTGAGTGCCAGTACCCCTTCTGATGAGAGACTTACTGACCAAGAGGAGGAGCTCA 2986  
Db  
2460 attcttltgaacaagagacagaggtgtctgtatatacaaaacccatcactgtgtg 2519

Db 2987 ATTCTTTTGAACAGGAGACAGAACTGTCTGTGAATACCAACCATACATCTGTG 3046  
Qy 2520 accctcaattagatgatatgatatatctcccaagtaaacctgataagat 2579  
Db 3047 ACCCTCCATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3106  
Qy 2580 atcaagtcctgtgtggaagcagttcccaagcctgaggagaaacccgaggaatgttg 2639  
Db 3107 ATACAGTCTGTGTGAGACAGTTTCCCAACCTGGCAGAAACCGGAGATGGTTG 3166  
Qy 2640 gaaagcaagttcccccaggtatctaactgaagccagtgctacacagatcagtc 2699  
Db 3167 GCAAGCAAGTTCCCCACAGGTTATTAACCTGAAGGCGCACTGACACAGTATCACTC 3226  
Qy 2700 agacagtaacccttgatgagagagcctgagaagaaattggccgcatgttcaagata 2759  
Db 3227 AGACAGTACCCCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3286  
Qy 2760 atcaacagggatcctcgtctgttccatcccttggaataactccctgtacacgtt 2819  
Db 3287 ATCCAAACAGGAGATCTGTGTTCTGTGTTCTGTGTTCTGTGTTCTGTGTTCTGT 3346  
Qy 2820 aggaagccttgagacaaatgattatcgacagtaacaggaacttgagagaggtataaaga 2879  
Db 3347 AGGAAGCTGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3406  
Qy 2880 gtgcagagacatacaacaaacagtcacgaacccctataaccccttgagagccctccgct 2939  
Db 3407 GTGCAGGACATACACCAACGAGTCCGCAACCTTATACCTTGTAGAGCCCTCCGCT 3466  
Qy 2940 gaaagaaacttgatagatgttggaacttaaaagatgcctctctcgtcgtgagataac 2999  
Db 3467 GAACGGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3526  
Qy 3000 cccactagcaaacacacttttgccttgatgagagagatccaagtaagggagaaacggg 3059  
Db 3527 CCACTACCAACCACTTTTGTCTTCAATGAGAGATCCAGGTACGGGAGAACCGGG 3586  
Qy 3060 caactacacttgacccagcctgcccagaaggttcaagaactcccgacatcttgagaa 3119  
Db 3587 CAGCTACCTGGAGCCGAGTCCGCAAGGTTCAAGAACTCCCGACATTTTGTAGAGA 3646  
Qy 3120 gacctacaagaggaacttgccaacttcaagataccaaacccctcagtgatcctccacg 3179  
Db 3647 GCCCTACACAGGAGCTGTGCGCAACCTCAAGATCCACACCTCAGTGACCTCTCCAG 3706  
Qy 3180 taactgagtagcctgcttctgtgagggagcccaacaagaagatcttagaaggtacgaag 3239  
Db 3707 TACGTGATGACCTGTCTGTGCGGAGCCACCAACAGAGACCTCTTAAAGAGTACGAG 3766  
Qy 3240 gcaactacacttgatgttgcacactagctacagagccttgcttaagaagccagatc 3299  
Db 3767 GCACACTGCTGGAATTTCTGACCTAGAGCTACAGAGCTCTGCTTAAAGAGGCCAAT 3826  
Qy 3300 tgcagagagagagtaacttgggttgcagatttgagggagggagcagatggtgagc 3359  
Db 3827 TGCAGGAGAGAGTAACTACTTGGGTGACGTTGGGGGGGGGCGAGCATGGCTGAGC 3886  
Qy 3360 gaggcagagaaagaaactgtatccagataccagcccaacacagccaaacaaagtga 3419  
Db 3887 GAGGACAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3946  
Qy 3420 gagtttttggagacagctggaatttgagagcttgatcccgaggttttgagacttagca 3479  
Db 3947 GAGTTTTTGGGACAGCTGGAATTTTGGAGACTGTGATCCCGGGTTTGGACCTTATGCA 4006  
Qy 3480 gcccaactctacccgcttaacaaagaaagagaggtatctccttgaggtccttgagacag 3539  
Db 4007 GCCCACTCTACCCGCTAACCAAGAAAGGGGGAATTTCTCTGGGCTCTGAGACACAG 4066  
Qy 3540 aagcatttgatgctatcaaaagcccttgagcagacactgcttgccctccctgac 3599  
Db 4067 AAGCAATTTGATGCTATCAAAAAGCCCTGAGGCGACACTGCTGCTGCTGCTGCTGAC 4126  
Qy 3600 gtaactaacaccttacccttctgtgagatgagcgttaaggaatagccgaagagttta 3659  
Db 4127 GTAACTAAACCTTTTACCCTTTTATGTGATGATGATGATGATGATGATGATGATGAT 4186  
Qy 3660 acccaaaccttagagacataggaagagacgttgcctacactgtcgaagaagctgtacct 3719  
Db 4187 ACCCAACCTTAGACCATGAGAGAGACCTGTTGCTTACTACTGCAAAAGAGCTTGTACT 4246  
Qy 3720 gtgagcaggtgtgtgcccgtatgtctggaagctatcgacagctgttgccatactgtcgaag 3779  
Db 4247 GTAGCCAGTGTGTGCGGCTGATGTGTGAAGGCTATGACAGCTGTGCTGCTGCTGCTG 4306  
Qy 3780 gacgtacacaaatlgacttggagacagataataacgttaataagcccatgcatggag 3839  
Db 4307 GACGCTGACAAATTTGATTTGGGACAGATATATAGTATATAGCCCCCATGATTTGGAG 4366  
Qy 3840 aacatgltcgcagccccagaccgaatggaatggaacagcccgcatgaaccactaaca 3899  
Db 4367 AACATGTTGGGAGCCCGGACAGCATGATGATGATGATGATGATGATGATGATGATGAT 4426  
Qy 3900 agcctgtcttcaagaagaggttcaacttgcctccacagccgctcacaacctgcact 3959  
Db 4427 AGCCTGCTTCTCACAGAGAGAGGTCACCTTTCGCTCACAGCCGCTCTCAACCTGCTC 4486  
Qy 3960 ctctgctgaagaagactgaatgaacagatgacatcagatctgacatcaactatgtatgag 4019  
Db 4487 CTCTGCTGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4546  
Qy 4020 gagaactgggttccgcgaagacacttacaagataccgctgactgagagaagltgctaaccttg 4079  
Db 4547 GAGACTGGGGTCCGCAAGGACCTTACAGACATCCCTGATGAGAAATGCAACTG 4606  
Qy 4080 ttacttgagaggaagcagatgtgtgtggaaggaagagatgtgtgagggcagtggtg 4139  
Db 4607 TTACTGACGGAAGAGCTATGTGTGTGGAAGGTAAGAGATGAGTGGGGGCGACAGTG 4666  
Qy 4140 gacgggacccgcagacttctgagcagcagccttgcggaaggaacttcaacgcaaaaggtc 4199  
Db 4667 GAGGGGACCGGACAGATCTGTGGCCAGACCTCCGGAAGGAATTCACAGGCAAAAGCT 4726  
Qy 4200 gagctatggccctcagcaagacttgcggttgccgaaggaatcaaaacttcat 4259  
Db 4727 GAGCTATGGCCCTCACAGCAAGCTTGGCGTGGCCGAAGGAATTCATTAACATTTAT 4786  
Qy 4260 aaggaagcagatgtacttctgtgactgtcacacgttcaaggggccaactcataaaagag 4319  
Db 4787 ACGGACAGCAGGTATCTTGTGCACTGACACGTACAGGGGGCCATTTATTAACAAAG 4846  
Qy 4320 ggtgttcttaacctcagcagggaggaataaagaacaagaaggaatcttaagcctata 4379  
Db 4847 GGGTTGCTTACCTCACAGGAGGAGGAATTAAGAACAAAGAGGAATTTTAAGCTATTA 4906  
Qy 4380 gaaagcttaacatttgcacaaaggtctagctatatatacactgtccttgcacatagaagcc 4439  
Db 4907 GAAGCCTTACATTTGCCAAAGAGGTAGCTATTATTAACACTGTGCTGAGATCAGAAAGCC 4966  
Qy 4440 aaagatctcatatctagaggaacagatgtctgacgggtgtgccaagcagccag 4499  
Db 4967 AAAGATCTCATATCTAGAGGAAACAGATGCTGACCGGTTGCCAAGAGGACACCCAG 5026  
Qy 4500 gctgttaacctctgctctaaatagaagcccaagcccgaaacccgaagaaagaaagac 4559  
Db 5027 GCTGTAACTTCTGCTTAAATAGAAAGCCCAAGGCCCAAGCCAGACGACAGATAC 5086  
Qy 4560 accctagaagacttgcaaaagataaaagaaatagaacagttctcgtagaactccgaggg 4619  
Db 5087 ACCCTGAAGACTGCGAAGAGATTAAGAAAGATAGACACTTCTGAGACTCCGAGGGG 5146  
Qy 4620 acctgctatacctcatalaggaagaaatcctgcccacaaagaaaggtttagaatagtc 4679  
Db 5147 ACCTGCTATACCTATATGGAAGAAATCTGCCCAAGAGAGGTTAGAAATATGTGC 5206







/translation="MGOTVTPPLSTIDHWTEYKSAHNLSNOVKKQPMOTFCVSEMP  
TFDVGWSEGEFENSELLAVKAILIFQGPSHPDQETITLFWDLADEPPWPKMLN  
KPRKPGRIALAEKKNHSAEKVPSPIHLEIEPPAMEPDSVPPYPAGAGK  
PSAPGAPAVEGPACTSRGATPERTDIALPELITGPPITPGOLOPLQWPPSS  
ADLYNMKTNPFSEDPORLTLGLVSELMFESHOPRQOOLLOTLTTEERKLTLEA  
KKNVSGADGRPTOLONEIDMGPLTRGCMQYNTAEGRSLKITYRQALVAGLSAR  
TNLAKEVWOGPNEPPSVLERMEKREPRPTDPDSFAOKASMLAATIGOSALDR  
KTLQRLGLEDAEHLADLVKEAEKVTIKRETBEEBROKREKEREERERKRRKREKNT  
KILAAVEGKSNRERERDFKIRSGPROGSLNRPPLDKOCAYCEKCKGMARDCK  
KNGKGLVLALEEDKD"  
2728. .6309  
/gene="pol"  
/note="complete pol ORF, gag-pol is expressed as a  
polyprotein"  
/codon\_start=1  
/product="polymerase"  
/protein\_id="CAC39617.1"  
/db\_xref="GI:14275839"  
/translation="GRGSDPLPEPRVTLKVEGQPVFVLDGKASHVLLQPKIKLD  
KKSWMGATGQOQYPTWTRRTVDLGVKRVHSFLVTECPAPLIGRDLTKMGAQISE  
BOGREVSANKRITVLQDLDEYRLVSPKPDONIQMLEPQAMETAGMGLA  
KQVPPVQOLKASAPSVRQYPLSKAREGIRHVQRLIQOGLIVVQSPWMTPLD  
VRKGTNDYRPVODLREYNKRKQVODIHPTVPYMLLALPDRSMYVLDLKDQFEL  
RLHPTSOPLAFEMWPDGAGRTGOLTPDKGNSPTTJEDALHRLDNFRLOHFO  
VTLLQYVDDLACATKQDCLEGTAKLLLELSDGYASAKKNOICREYVYGSIR  
GGOWMLEARRRYVQIPAPTTAKOVEFEETADFCKMLIEGTALAPLPLTKEG  
EFSWAPBEHQKAFDAIKRALLSAPALAPDVIRLDEKRGKARVAVDTQTLKMPR  
PVAYSKLDIVASGMPICLKAIAVALIVADAKTLGQNTIYIAHALENTVROP  
DRMWTNMTHTYOSLLTERVTFAPALNATLPEETDEPRLIMASDEGSAQAEIM  
KDLDITPLGMLTQFDDGSSYWGKMGAGAAYVDGRTIMASDEGSAQAEIM  
ALTOALRAEKSINTYDSRATPAVHAITKOGILLASAREKKEKKEISLE  
ALHPRKRLAIIHCPEHQKAPISRGONADPRAVKAQANQNLPHIETKRAPRGO  
YTLDEMOEIKRIDFSETEGCTYSDEKELIPHEGLEIVQOJHRLHIGLKLQOL  
VLTSEYHVLRLPVADSVYKHCVPQOLNANPRLPEPKRLGSHPGAHEWDETEVK  
PAKYGNKYLVEVDFSGWEAYPKKSTVYVAKLIEETFPFGPIKVIYSGONGPA  
FVAOVSQGLAKLIGDKMLHCAYRPOSGOVERMNRRTKETLTLTETGIDMMAL  
PVLVFRVNTPGQGLTPEYLLYGSPPLAEIAFAHSDVLLSOPLSRLKALMVRQ  
RAMQOLREAYSGDQVPHRFQVGSYVRRHRGNLETPRKSPYVILLTTPAVKE  
GTPVTHSHVAPPPDSGMAKTEPNLKLRLHRYVPYVNNSS"  
6185. .8149  
/gene="env"  
6185. .8149  
/gene="env"  
/gene="env"  
/codon\_start=1  
/product="envelope"  
/protein\_id="CAC39618.1"  
/db\_xref="GI:14275840"  
/translation="MHPILSRHRLPIRGKPKRLKIPLSFASIAWFLTLSTPQNGK  
RLVDSNHSKPLSLWLLTSGTGININSTOGEALGTWMPLELYCLSVIPGLNDQA  
PPVILRAYGFCVCPGNNEVEYCGNNGKMCSTTSNGKMKWVSOODRYS  
FYNNPTSTNOFTYGHGRKMDQORQOKVYRNKOISCHSLDLYLKISTPEGKQONTO  
KWNNGISWGLVYGGSGRKGSVLITRIETOMPEPAIIGNKLAOGCPIDORP  
SNPDSYNTSGSVPEPNTITKTAKLSLLOGAFOALNSTPATSSCHLNTSE  
PYEGMARGKENVKREHDOCTGWSOKLTLVEGSGCTGAMPBPHQMLCANTGA  
FNRTSEGYLVGYDRMACNTGTLPGCATVFNQTKFCVQVQJLVPPVYVYKEAVL  
DEVDYRNKRPKREPSITLAVMLGLGAVAGVGTGAALITGQOLEKESINHLVTE  
DIOALEKSYSNLEESTLSSEYVIONRGLDILFLKEGIGCALKECCCFYDHSAL  
KDSMSKRLERLEKREKREKAGGMEGPNKSPWTTLSALTGDLVILLLLTYGPC  
LINRFAVREYSAVRIMVLRQYOGIPS"  
8216. .8918  
8216. .8918  
LTR  
polyA\_signal  
BASE COUNT 2435 a 2174 c 2253 g 2056 t  
ORIGIN

Query Match 79.1%; Score 6493.2; DB 14; Length 8918;  
Best Local Similarity 87.9%; Pred. No. 0;  
Matches 7332; Conservative 0; Mismatches 798; Indels 210; Gaps 15;

OY 1 gtggtgaagcgtgtggcccgagcggtctggaataaatacctctgtgttgatc 60  
|||||

Db 562 GTGGCGTACGACTGTGGGCCCCAGCGCGCTCGGAATAAAAATCTCTGCTGTTGCATC 621  
OY 61 aagacgcctctcgtgaagatctgggggtctgcctctccgagcccgagagggagat 120  
|||||  
Db 622 AAGACCCCTTCGTGAGTATTTGGGAGTGTCCCTCTTCGAGTACAGACGAGAGGAT 681  
OY 121 tgtctcttaagcctctcaatttctgtctgttgcggagaaatctctgcagaccctta 180  
|||||  
Db 682 TTTAACTGACATGCGCTTTCAGTTTGGTGTGGCGGAAACCGCGCATCCCTCA 741  
OY 181 caccgagaaacgacttgagagtaaaaggaatccccctttgaaagt----gtgtgtgtc 236  
Db 742 CACCCGAGAACCGACTTGGAGGTAAAGGATCCCTTTGAACTGTGAGTGTGTGTGTC 801  
OY 237 ggcgcgct 296  
Db 802 GGGCGGCGTCT 861  
OY 297 tcccttaagacgcgtaagacttgagagcttgatcaagagaagctgtctagaagatacaag 356  
Db 862 TCCTCTCAGACCGTAAGGACCTGGGAGCTGTGATCAGACAGCTGTAGAGGATACAG 921  
OY 357 gctgcacaccttgaggagcccccggaggagtgaggagagcgaagagacccctggtgtctc 416  
Db 922 GCTCCACCCCTGGGGGAGCGCCCGGGAGGTGGGAGAGCCAGAGCCCTGTGTGTCTCC 981  
OY 417 tactctcgtcagaagacgcgactctgtctgttgaagcgaaagcttccccctccgagcg 476  
Db 982 TTTCTGTGCTGCTGAGAGACCGAGCTGTGTGTAAGGAAAGCTTCCCGCCGCGCGC 1041  
OY 477 tccgactcttctgcgtctgttgaagaacgagagcggtctgtgtctgtgtctgtt 536  
Db 1042 TCCGACTCTTTCCTGCTGTGTGGAAGACGGGAGCGGCTGCGGTGTGTGAGACTTGTG 1101  
OY 537 gttctcgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 596  
Db 1102 GTTCTGT 1161  
OY 597 aacagtacaccccccctagtttgcactctgcacatctgagactgagatagatccaggct 656  
Db 1162 ACGGTGACGACCCCTCTTATTTGACCTGACCATTTGACCTGAAATTCAGAGGCT 1221  
OY 657 cataatlttgatcagtttaagaagagagacttggagactcttctgtcctctgaatg 716  
Db 1222 CATATTTGTGATGTCAGTTTGAAGAGGAGACTTGGAGACTTCTGTGTGTGATGG 1281  
OY 717 ccaacattcgaatgttgaatgacatcaagaagagacatttaattctgaatattctgtct 776  
Db 1282 CGGACATTCGATGTTGATGGCCATCAGAGGGAGACTTATTTCTGAGATTAATCTGGCT 1341  
OY 777 gtaaggaatcaatttcaagacttgagacgcgctctcaactgatalcaagagccctatatc 836  
Db 1342 GTTAAAGCAATTAATTTTACAGACTGACCCGCTCTCATCCGATCAGACGCCCTATATTC 1401  
OY 837 cttaagtgcaagatttggcagaagaatctctcgcaatgggttaaaccaatggctaaataa 896  
Db 1402 CTTACGTGGCAGATTGGCAGAGATCTCCGCCATGGGTTAACTTTGGCTGAATAG 1461  
OY 897 ccaagaaagcagtcgcccgaaatctctgtctcttgaagaaacaaacactcggcgaa 956  
Db 1462 CCAAGAAAGCCAGTCCCGCAATTCGCTCTGTGAAGAAACAAACACATCGCGTGA 1521  
OY 957 aaagtcgagccctctctcgtatcaaccgagatcaagagagcccgacttggcgga 1016  
Db 1522 AAAGTCAAGCCCTC--TCCATATATCTACCCAGATTAAGAGCGCGCGCTTGGCCGA 1580  
OY 1017 acccaacactgttcccccaccccttaccacacagagtgctgtgaagaga--cctctgc 1075  
Db 1581 ACCCCAATGTGTCCCAACCCCTTATTCGACACAGGAGTGTGAGAGGACCTTGTGC 1640  
OY 1076 ccctctgagctcgtgtgtgtgaagagacttgcgagactcgagacggagagagcg 1135  
|||||  
Db 1641 CCTCTCTGAGGTCCGGCGGTGGAGGAGACTCTGACAGGACTCGAGACCGGAGGCGC 1700  
|||||

QY 1136 caccccgagcgagacagagatcgataltacagctgcacactatggcccccatt 1195  
|||||  
Db 1701 CACCCCGAGCGAGCAGACAGATGCGACATTTACCGCTGGCAGATAGCCCTCCCAT 1760  
|||||  
QY 1186 gccaggggccaattgcagccccctcagatattgccccttctctctcagatctctataa 1255  
|||||  
Db 1761 ACCGGGGGCCAATTGACGCCCTCCACATTTGGCCCTTTCTCTGACATCTCTATAA 1820  
|||||  
QY 1256 ttggaataactaacatcccccttctcgagagatcccccaagcctcaggggttggtga 1315  
|||||  
Db 1821 TTGGAAACTAACCATCCCTCTTCTCGAGAGATCCCAACGCCCTCAGGGGTTGGTGA 1880  
|||||  
QY 1316 gtccctatgtctctcaccagcctacttggatgatgtcacaagctgtcgcagacact 1375  
|||||  
Db 1881 GTCCCTTATGTTCTCTACACGCTACTTGGGATGATGTCAACAGCTGCTGCGACACT 1940  
|||||  
QY 1376 cttaacaacgagagagcgagagaatctcgtctagaagcttagaaaaatgttccctgggc 1435  
|||||  
Db 1941 CTTCAACACGAGAGCGCAAGAAATTTGTTAGAGCTAGAAAAATGTCTCTGGGCGC 2000  
|||||  
QY 1436 cgaacggcgacccacgcagctgcaaatgagattgacatggatcttcccttgactggcc 1495  
|||||  
Db 2001 CGACGGGGGACCCACGCGAGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTGCC 2060  
|||||  
QY 1496 cgttgggaactacaacacgcgtgaagtgaggagagcttgaataatctatcgccagctc 1555  
|||||  
Db 2061 CGGTTGGGACTACAAACGCGCTGAAGGTAGGAGACCTTGAATAATCTATCGCCAGGCTCT 2120  
|||||  
QY 1556 ggtggcggtctcgggggcgctcaagaagcccaactaatttggctlaagtgagaagt 1615  
|||||  
Db 2121 GGTGGCGGCTCTCCGGGGCGCTCAAGACGGCCACTACTTGTGGCTAAGGTAGAGAGGT 2180  
|||||  
QY 1616 gatcagggagacgaacacactccctcgtgatatctctgaaggtctatgaagccttcag 1675  
|||||  
Db 2181 GATCAGAGGACCGAATGACCTCCCTCAATTTTCTTGAAGGCTATGGAAGCCTTTCAG 2240  
|||||  
QY 1676 gcggttacccttctgtatcctactcaagaagcccaagaaagctcagttggccctggcct 1735  
|||||  
Db 2241 GCGGTTACCCCTTTTATCTTATCTACCTCGAGGCTCAAGAAAGCCTCAATGCTCTGGCCT 2200  
|||||  
QY 1736 catgggcagctcgctctgtgatatcagaagaactctcagaagcttgaaggttlaagaa 1795  
|||||  
Db 2301 CATAGGACACTACCCCTGGATATTCAGAAAGAGCTTTCAGAGCTGGAAGGCTTACAGGA 2360  
|||||  
QY 1796 ggcctgagttacgtgatctagtgagagagcagaagaagtgatattacagaagggagacaga 1855  
|||||  
Db 2361 GGCTGAGTTACATGATCTAGTGAAGGAGCGAGAAAGTGTATTACAAAGGAGACACAA 2420  
|||||  
QY 1856 agagggagaagacagagaaaaaagaagagagagagaagaagggagaaagcgtgataag 1915  
|||||  
Db 2421 AGAAGAAAGGAGCAAAAGAAAGAGAGAAAGAGAAAGGAGGAGAAAGCTAATTA 2480  
|||||  
QY 1916 acggcagaagaagaatttgaactaagaacttgcgcagcagtggttgaagggagaagcgag 1975  
|||||  
Db 2481 ACGGCAAGAGAAAGATTGACTAAGATCTTGGCTGCAAGTGCTTGAAGGAGAAAGCAATAG 2540  
|||||  
QY 1976 gggagagagagagatlttagaanaaattagtcagagccctagacagctcagggaaactggg 2035  
|||||  
Db 2541 GGAAGAGAGAGAGATTAGGAAATTTAGTCAGGCGCTAGACAGTCAAGGAACTCTGGG 2600  
|||||  
QY 2036 caatagagacccactcgaagaagacagagtgctgatatgttaagaaaaagagcaactgggc 2095  
|||||  
Db 2601 CAATAGAGACCCCACTCGAAGAGACCAATGTGCTATTTTAAAGAAAAAGGACACTGGGC 2660  
|||||  
QY 2096 aaggaacttgcaccaagaagaagaacaagaagacagagtgctcagctcctagaagaagataa 2155  
|||||  
Db 2661 AAGGAGACTCCCAAGAAAGGAAACAAAGAGTGAAGGTCTTACGCTCTGGAAGAAAGATTA 2720  
|||||  
QY 2156 agattagggagagaggggttgcgaaccctcccgagccagggtaactttgaaggttga 2215  
|||||  
Db 2721 AAGACTAGGAGAGCGGGTTTCGACCCCTCCCGAGCCAGGCTTAACCTTTAAAGGTGA 2780  
|||||

QY 2216 ggggcaacagttgagttcctggttgataccggaagggagcattcagttcctataacc 2275  
|||||  
Db 2781 GGGGCAACCAAGTTGAGTTCTGTTGATATACCGGAGGAAACATTTCACTACTACAGCC 2840  
|||||  
QY 2276 attagaaacataaagaanaaaatccctggggtgagtggtgtccacagggcagaagcagta 2335  
|||||  
Db 2841 ATTAGGAACCTAAAGATTAATAATCTGTTGGTGTGATGGGTGCCACAGGGCCAAACAAATTA 2900  
|||||  
QY 2336 tcaatgtaactaccgaagaacgcttgaacttggagtggtggaacgggtlaaccactgttct 2395  
|||||  
Db 2901 TCCATGGACACTACCCGAAACACAGTTGAGTGGAGTGGGAGCGGATTAACCCACTGTTCT 2960  
|||||  
QY 2396 ggtcaccctcagtggtgcacagtaacccttctagtgaaagcttacttgaacaaatggagac 2455  
|||||  
Db 2961 GGTGATACCTGAGTGGCCGACACCCCTCTTAGGTAAAGACTTACTTACCAATGAGGAGAC 3020  
|||||  
QY 2456 tcaattctcttgaacaagaagaacgaagaagtgtctggaataaacaacccatacgt 2515  
|||||  
Db 3021 ACAATTTCTTTTGAACAAAGGGAACCAAGAGTGTGCAATTAACCAACTATCACTGT 3080  
|||||  
QY 2516 gttgacccctccaattagatgataatactgaactatcttcccaagtaaacctgtatca 2575  
|||||  
Db 3081 GTTGACCCCTCCAAATTAGATGAGAGCAATATGACTATATTTCTCCCTAGTAAACCTGATCA 3140  
|||||  
QY 2576 agatataagtcctggttggagcagtttcccaagccttgggagaaacgcagagatggg 2635  
|||||  
Db 3141 AATATATCAATTTCTGTTGGAACAGTTTCCCAAGCCTTGAGGCAACCGCAGGATGGG 3200  
|||||  
QY 2636 ttggcaaaagaagttcccccacaggttattcaactgaagggcagtgctacacagttac 2695  
|||||  
Db 3201 TTTGGCAAAAGCAAGTTTCCCCACAGGTTATTAACATGAAGGCAATGCTGTCACAGCTTC 3260  
|||||  
QY 2696 agtcagacagtgatcccttggatgagagaggtctcgaagaagatttggccgcatgttcaag 2755  
|||||  
Db 3261 AGTCAGACAGTACCCCTTGAGTAAAGAGCTCGAAGAAATTCGCGCGCATGTTCAAG 3320  
|||||  
QY 2756 attaatcaacaagggcaltccctagctcgttcccaactcccttggaaatactccctgtacc 2815  
|||||  
Db 3321 ATTAAATCAACAGGGGCAATCTAGTTCCTGCTCAATCTCCCTGGAATTTCCCTGCTAC 3380  
|||||  
QY 2816 ggttgaagaagccttggacccaatgatalatcgaacagtaacagagcttgaagaagtcataa 2875  
|||||  
Db 3381 GGTTAGAAAGCCTGGGACTAATGACTATGACACAGTACAGAGCTGAGAGAGTCAATTA 3440  
|||||  
QY 2876 aaggttgcagagacataacccaaggttccgaacccctataactctttagcgcctccc 2935  
|||||  
Db 3441 ACGGTTGACGATATACACCAACAGTCCGAAACCTTATTAACCTTGTGCTCC 3500  
|||||  
QY 2936 gcttgaacggaacttgaacagatattgacttaaaagatccttcttctcctgagatt 2995  
|||||  
Db 3501 ACCCCACGAGAGCTGTATACAGTATTTGACTTAAAGATCCCTTCTTCTGCTGAGACT 3560  
|||||  
QY 2996 acaaccactagccaacacttlttgccttcgaatggaagataccaggtlacgggaagaac 3055  
|||||  
Db 3561 ACATCCCATGAGCCACACACTTTTTCCTTGAATGAGAGATCCAGGTGGGAGAAAGAC 3620  
|||||  
QY 3056 cgggcagctcaacttgcagccgacgtgcgccaaggggttcaagaagctcccccagcatcttga 3115  
|||||  
Db 3621 CGGGAGGCTCACTTGGATGACGTGACCTGCCCAAGGGTTAAAAAACTCCCGGACACTTTTGA 3680  
|||||  
QY 3116 cgaagccctacacagggagccttgcgaacttcaagatccacaacacccctcaagttgacccctc 3175  
|||||  
Db 3681 CGAAGCCCTTACACAGAGACCTGGCCAACTTCAGATCCAAACACCCAGGAGAACCTCTCT 3740  
|||||  
QY 3176 ccagtagctgatactgcttcttgcgggagagccacaacaggaactgtctagaagttac 3235  
|||||  
Db 3741 CCAGTAGCTGATGATACCTGCTTCTGGGGAGGAGCCACCAACAGAGACTGCTTAGAAGGTAC 3800  
|||||  
QY 3236 gaagcaactatgtctggaatgtctgaacccctaggttaagagccttctcgttaagaaggcca 3295  
|||||  
Db 3801 GAAGGCACTACTGCTGGAATTTGTGACCTAGGCTACAGAGCCCTCCCTTAAGAGGCCCA 3860  
|||||  
QY 3296 gattgcagagagaggaaglaacatacttgggttgcagtttgcgggagggcgagcagtgct 3355  
|||||

|||||  
Db 3861 GATTTCAGAGAGAGAGTAACATCTTGGGCTACCTTTGGGGGGGGGAGGATGGCT 3920  
Qy 3356 gaaggaagcaggaagaaatctgtagtccagataccgcccacacacgaacgaagt 3415  
|||||  
Db 3921 GAGGAGAGGACGAGAGAGATCTAGTCCAGATACCGGCCCCCAACACAGCCAAACAGT 3980  
Qy 3416 gaagaagtttttgggagacagctgatttggcagacgttgatcccgagggttggagcctt 3475  
Db 3981 GAGAGATTTTTTGGGACACCTGGATTGTCAGACTGTGGATCCGGGTTTGGCAGCTT 4040  
Qy 3476 agagagcccaactaccggttaacccaagaaagggagatctctcgtggctcctgaaga 3535  
Db 4041 AGCAGCCCACTCTACCCATTAACCAAGAAAAGGGGAATTCTCGGGCTCTTAGCA 4100  
Qy 3536 ccagaagacatctgtagtctaaagagccctgtcgtgacgcacgtctcgtccctcc 3595  
Db 4101 CCAGAAAGCATTTGATGCTATCAAAAAAGCCCTGTGAGCGCACCTGTGCGCCCTCCC 4160  
Qy 3596 tgacgtacataaaccttaccctttagtggatgagcgtgaagggagtagccggagagt 3655  
Db 4161 TGAGGTACCTPAACCTTTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGT 4220  
Qy 3656 tttaaccacaacccttagacatgagagagacgtgtgctacactgtcaagaagacttga 3715  
Db 4221 TTTAACCCAAACTCTAGGACCATGAGAGACCTGTTGCTACCTGTCAAGAAAGCTCGA 4280  
Qy 3716 tccctgtagcagtagtgtgcccgtatgtctgaaggtatccgagcgtgtgcaactggt 3775  
Db 4281 TCCTGTAGCCAGTGTGTGGCCCATATGCTGAAGGGCTATGCGAGCTGTGCGCATACGCT 4340  
Qy 3776 caaggaagcgtgacaaattgacttggagacagaataactgtaataagcccccatgatt 3835  
Db 4341 CAAAGAGCGTGACAAATTTGATTTGGAGACAAATATACCTGTATTAACCCCTCATGCTT 4400  
Qy 3836 ggaagaacatgctcgcagagcccccagacagctgtagtaccacagcccgca tgaaccacta 3895  
Db 4401 GGAAACATCGTTGGGAGCGCCCGACGACGATGATGACCAAGCCCGCATGACCCACTA 4460  
Qy 3896 tcaaaagcgtcttctcaagaagagtgtaacttctgcctcaacagccgcttcaaccctgc 3955  
Db 4461 TCAAAAGCCTGCTCTCAGAGAGAGGTGACGTTGCTCCACACCCCTCTCAACCCCTGC 4520  
Qy 3956 caccctctgctgaagaagactgtagaacagtgactcaatgattgcatcaactatgat 4015  
Db 4521 CACTCTTGCTTAAGAGACTGATGAACCAAGTACTCATGATTGGCCATCACTATTGAT 4580  
Qy 4016 tgaagaagactgggttccgcaagagacttacagacataccgctgacttgaagaagtgtaac 4075  
Db 4581 TGAGAGAGCTGGGGTCCGCAAGAGACCTTACAGACATACCGCTGACTGAGAAATGTTAAC 4640  
Qy 4076 ctggttcaactgaacgaagacgtatgtgtgtgaagtaagaagtgctgtggcgagct 4135  
Db 4641 CTGTTCACTAGCAGGAACACCTATATGTGGAAGGTAAAGGATGTGCTGGGGGGGGGT 4700  
Qy 4136 ggttgaagcggagcccgcaagactctggccagagacgtccggaagaagaacttaagcgcaaa 4195  
Db 4701 GGTGAGCGGAGCCGACGATCTGGGCGACGACGCTGCGGAAGGAATCTTACGACAAAA 4760  
Qy 4196 ggttaagctcctgcccccaagcaagcttggcgctggccggaaggaatccataaacat 4255  
Db 4761 GGCTAGCTCAGTGGCCCTCACAGCAAGCTTGGGCTGGCCGAAGGGAATCATTAACAT 4820  
Qy 4256 ttataaggaacagagtagtcccttggactgcacaaagtatacggggcactataaaca 4315  
Db 4821 TTATACAGACAGCAGATGTGCTTTGGCAGTGCACACCTTACACGGGGCATCTATAAGCA 4880  
Qy 4316 aaggaggtgtcttaccgaagcaggaaggaataaagaacaagaaggaatctcaagcct 4375  
Db 4881 AAGGGGGTTGCTTACCTTAGCAGGAGGAGGAATTAAGAACAAAGAGAAATTTCAAGCT 4940  
Qy 4376 atlaagaacctacatttgcacaaaaggctagctattatacagtgctcgtgacatcaaga 4435  
|||||

---

Db 4941 ATTAGAACCTTACATTTTGCCAAAAAGGCTAGCATTTATACATGCTGCTGCATCAGAA 5000  
Qy 4436 agccaagatcctcatatctagaaggaaccagatggtaccgggttgcacagcagcagc 4495  
Db 5001 AGCCAAAGATCCCATATCTCAGAGGAGAACACAGATGGCTACCGGGTTGCCAAGGAGCAGC 5060  
Qy 4496 ccaggctgtlaaccttctgctataatagaacgcccccaagcccccaagaaccagacga 4555  
Db 5061 CCAAGGTGTAACTTTCGCTATGATGAAAAACCCCAAGGCCGAGAACCCGGACGACA 5120  
Qy 4556 gtacacctagaagactggaagagataaaaaagtagaacagttcttgagactcgga 4615  
Db 5121 GTACACCCCTAGAAAGCTGGCAAGAAATAAAAAGATPAACAGATTCCTGTGAACCTCGGA 5180  
Qy 4616 gggagacgtctatactcaatggaaggaatctctgcccccaagaaggttagaata 4675  
Db 5181 GGGAGACTCTTATCTCAGATGGAGAGAAATCTGCCCCCAAGAAAGGGTTAGAAATA 5240  
Qy 4676 tgtccacagatatacgtctcaaccacactaggaactaaacactgcagcagttgttag 4735  
Db 5241 TGTCACAGATATACATCGTCTAACCCACTAGGAACCTAACACCTGCAGCATTTGCTAG 5300  
Qy 4736 aacatccctatcatgttctgaagctaccagagatggtgtgctggtgtgaacattg 4795  
Db 5301 AACATCCCTTATCATGTTCTGAGGCTACAGAGAGTGTGACTGTGCTGAGTGAACATTTG 5360  
Qy 4796 tgtgcccctgcagcgtgtgtaatgtaactcctccagaatacctccagaaagagataag 4855  
Db 5361 TGTGCCCTCCAGCTGTGTTAATGCTATCTTCCAGAAATACCTCCAGAAAGAGCTTAAG 5420  
Qy 4856 gggaaagccacccagcgtcactgtggaagtgtactcaactgaggtlaaagcggctaata 4915  
Db 5421 GGGAAACCCACCCAGGCGCTCACCTGGGAAAGTGAACCTGAGTGAACCCCGCTAAATA 5480  
Qy 4916 cggaaacaatatctattgtttttagaacctttagaagatgtagaagctatacc 4975  
Db 5481 CGGAAACCAATATCTATTTGTTTGTGAGACCTTTTGTAGAGTGGTGAAGCTTTATTC 5540  
Qy 4976 tactaagaagagacttcaacgctgtgtgctlaagaanaatacgtggaagaaattttccaag 5035  
Db 5541 TACTAAGAAAGAGACTTCAACCGTGTGCTTAAGAAATATCTGAGGAATTTTCCAG 5600  
Qy 5036 atttgaatacctaaggtataggttcagacaatgtgtccagcttctgttgcaggttag 5095  
Db 5601 ATTTGAATATCACTAAGGTATAGGGTCAGACAAATGCTCCAGCTTTCGTCAGAGTAAG 5660  
Qy 5096 tcaaggacgtggccaagatatgtgggttattggaacgcaatttgcatacagaaccaca 5155  
Db 5661 TCAGGGACTGGCCAAATATTTGGGATGATTGGAACCTGCAATGTGCTATACAGACCCCA 5720  
Qy 5156 aagctcaggaagtagagagagatgaatagaacatlaagaagaccccttaccaaattgac 5215  
Db 5721 AACCTCAGAGCAGGTAGAGAGATGATGAATGAACATTAAAGAGACCTTACCAAAATTGAC 5780  
Qy 5216 cacagagactgcatlaatgattgtagtgcctcctgcctcctgtgttctttaggttag 5275  
Db 5781 CACAGAGACTGGCATTAATGATTGATGTGCTCCTGCCCTTGTCTTTTAAAGGTGAG 5840  
Qy 5276 gaacacccctggagcagtttgggtgtgagcccccataatgattgtcttaaggggagaccccc 5335  
Db 5841 GAACACCCCTGGAGAGTTTGGGCTGACCCCTATGAAATTTGCTTACGGGGAGACCCCCCC 5900  
Qy 5336 gttgacagaatctgaccttgcacatgctgtgactgtgcttcccaagccttgtctc 5395  
Db 5901 GTTGGCAGAAATTTGCTTTCACATATGTGCTATGTGCTTTCACAGCTTGTGTTCTC 5960  
Qy 5396 taagctcaagcgtcagatgggtgtgagcgaagcgttggaaagcagctccgggaagccta 5455  
Db 5961 TAGGCTCAAGGCGCTGAGTGGGTGAGGCGCACAGGCTGGAACACCTCCGGAGGCTTA 6020  
Qy 5456 ctaaggagagacttgaatttccacatgcttccaaagtgtgagatctatgttag 5515  
Db 6021 CTACAGAGGAGACTTGCAGATTCCACATGCTTCCAAAGTTGGAGATTCAAGTATGTTAG 6080  
|||||

Qy	5516	acgcacacgctgaggaagaacccctcagaaactcgtgtgaaagggaacctatctcgtactcttgac	5575
Db	6081	ACGCCACCGTGGAGGAACCTCGAGACTCGGTGGAAAGGAGACCTTATCTCGTACTTTGAC	6140
Qy	5576	caacacacgagctgtgaaagtcgaagaatccccacctcgtgatacctatcccaagttaa	5635
Db	6141	CACACCAACGGCTGTGAAAGTCGAAGGAATCCCACTGGATCCATGGATCCACGTTAA	6200
Qy	5636	gcgcgcgcacccctccgcagctcggggtgagagccgaaaaagactgaaatcccccctaact	5695
Db	6201	GCCCGCGCCACCTCCCGATTCGGGGTGGAAAGCCGAAAGACTGAAATCCCTTAAGCT	6260
Qy	5696	tcgcctccatcgctcgtgtctctactcttaacaataactccccagccagtgtaaagcc	5755
Db	6261	TCGCCTCCATCGCGGTGTTCTTACTCTGTCAATTAACCTCTCAATGATGATTAACCC	6320
Qy	5756	ttatagacagctcgaaaccccatagaaaccttatccctactcgtcgtatattgaacctg	5815
Db	6321	TTGTGGACAGCCCAACCTCCATTAACCCCTTATCTCTCACCTGGTTACTTACTGACTCG	6380
Qy	5816	atagaggtgtgactgtaataatgcactcogaagtggtgtgtcctcagaagccactgtgtgacctg	5875
Db	6381	GTAAGGTATTAAATTAATTAACAGCACTCAAGGGGAGGCTCCCTTGGGACCTGTGGCTCG	6440
Qy	5876	aactgcattctcgtcccgatltgaatlaacccgcgtgtta-----aaagccacactccca	5929
Db	6441	AATTAATATGTGCTGCTTCGATCACTAATTCCTGTCTCAATGACACGAGCCACCCCTCG	6500
Qy	5930	accgaagccgtgaagtataagggtgtcattcgtgtgccagg----cacagagaagaagaatact	5986
Db	6501	ATGTACTTCGGCTGTACGGGTTTATCGTTTTCGCCAGAGACCCCAATAATGAAATATAT	6560
Qy	5987	gtgagggtctctgaggaaatccctcgtcgtagagaatgagctcgtcactcccaagatbag	6046
Db	6561	GTCGAATTCCTCAGGATTTCTTTTGGCAAGCAATGAGCTGCAATCTTCTTAATGATGGGA	6620
Qy	6047	acgtgaaatgccgatctctctccagaagccgggtaaaatctccttctgt-----	6095
Db	6621	ATTGGAATATGGCAGCTCTCGACAGACAGAGTAAGTTACTCTTTTGTAAACAACTCTA	6680
Qy	6096	-----caattccggcccgggcgaagtac	6117
Db	6681	CCAGTTTATATCAATTAAATTATGCGCATGGAGATGGAATTTGGCAACAGCGGGTAC	6740
Qy	6118	aaaaatgtaactataataagaataagagctgtctcccatcagaactagatattctaag	6177
Db	6741	AAAAAGATG-TAGGAATAATGACAAATTAAGCTGTCTCATTCGTTAGACCTAGATTACTTAA	6799
Qy	6178	ataagtttcaactg--aaagaaacacggaaataatattcaaaagtgtgataatgtatgagc	6235
Db	6800	ATTAAGTTTCACTAAAAAGGAAACAAAGAAATATTTCAAAAGTGGGTAATGTGTATATCT	6859
Qy	6236	tgaggaaatagtttttat-----catactgcggggagcagaggtcccaacttaaacactcgc	6291
Db	6860	TGGGGAATATGTACTATATGAGAGCTCTGGGAGAAAGAAAGATCTGTTCGACTATTTCCG	6319
Qy	6292	cttagatctgagagcgggagcagaagaacccctcgtgtgcaatgaggaccgataaagtactgact	6351
Db	6920	CTCAGATATGAACCTAGATGAGATCGGTTTGTATAGGACCAATTAAGGGTTTGGCC	6379
Qy	6352	gaacagggagcccgccgctcgtgagccacgcgactaactgtccggtgcccaatlaactcgt	6411
Db	6980	GAACAAGACGCTCCATTCACAGA-----A	7003
Qy	6412	ctgcggcctgacataacacagccgcctagcaacagtaaccactggtatgttctctacaac	6471
Db	7004	CAGAGGCCATCTCTTAACCCCTGTGATTACATTAACAACCTCTGGATCCCTCCCACTGA-	7062
Qy	6472	acgcctagaactcccaggtgtctcgttlaagacagagacagactcttcagttcctac	6531
Db	7063	-----GCTTAACATCACTATTAAACAGGGGCGAAACTTTTGGCTCANC	7108
Qy	6532	caggagagcttcccaagccactcaactccacccgagcccttgatgacactctctctgttgct	6595
Db	7109	CAGGAGGCTTTCAAGCTTTTAACCTCCAGACTCCAGAGGCTTAACCTTTCTGTGTGGCT	7166
Qy	6592	tgcttatccctcagggcctccctctattatgagggaatgagctaaagaaataatctcatgt	6655
Db	7169	TGCTTAGCTTCGGGCCCACTTACTATGAGGGAATGGGTAGAGGAGGAAATTCATGTG	7222
Qy	6652	accaagagcatlagaaatcaatgatacatgaggggtcccgaaataaacttaactcctctgaa	6711
Db	7229	ACAAAGAACATTAAGACCAATGTACATGGGATCCCAAAATTAACCTTACTCTGAG	7286
Qy	6712	gttccgggaaggggacatgcaatgaagaaagctcccccatcccaacacacacttgcata	6777
Db	7289	GTTTCTGAAAAAGCAGCCTGCAATAGGATGGTTCCCATCCCAACACTTTGTATAC	7344
Qy	6772	agtaactggttttltgagcaggccctcagaaaaactgatttaagtactcgtgtataacagg	6831
Db	7349	CACACTGAAGCCTTTAAATGAAACCTCTGAGAGTCATATCTGTGATACCTGTTATACAG	7408
Qy	6832	tggtggagcatgcaatactcgtgtlaacccccgtgttctccactcagttcttaacaaatcc	6891
Db	7409	TGTTGGGCAATGTAATACTGATGATTAACCTTGTTTCCACTTGTGTTTCAACCAACT	7466
Qy	6892	aaagattctgtgtcatggtgtccaaatcgtgtcccgagtgtaactacatccctgaggaagt	6951
Db	7469	AAAGACTTTTGGTTATGTGCTCAAAATTTCCCGGGGTACTACTATCCGAAAAAGCA	7528
Qy	6952	gtccttgatgatatgactacgtgtataaacccgaacaaagaagaacggtatcccttacc	7011
Db	7529	GTCTTGATGAATGTACTATAGATATATATGCGCCAAAAAGAGCCATATCCCTTGACA	7588
Qy	7012	ctagctgtaatgctcggatattagggaacggccgtgtgctglaagaaacaggacgctgcctg	7071
Db	7589	CTAGCTGTAATGCTCGGATTTGGGATGGGCTGCAAGCTGGGGAACAGAACCGCTGCCCTA	7648
Qy	7072	atcacagagcacacagcagctlagagaagaagactgtgtagctacactggtccatgacgaa	7131
Db	7649	ATCACAGACCGCACAGCAGCTGGAGAAAGACTTAGTAACCTACATCAATTTGTAAACGAA	7708
Qy	7132	gattccgagcctaaagagatcgtgttagcgaacctlaagaaagctccgacttctgtct	7191
Db	7709	GATCTCCAAACCCCTAGAAATATCTGTCAGTAACCTGAGGAATCCCTAACCTCTTATCT	7768
Qy	7192	gaagtggttctaaagaacggagggatattgactcgtgttcttaagaagaagtgggtta	7251
Db	7769	GAGTGTGTTTACGAACAAGAGGGGTTGATCTGTATTATTTCTTAAAGAAAGAGGGGTTA	7828
Qy	7252	tgctgagccttlaaagaagaatgtgtctctatgtaatacctcaacgaagacatcaagagac	7311
Db	7829	TGTGTAGCCTTAAAGAGGAATCCTGCTTTTATGTGATCATTCAGAGCTATACAGGAC	7888
Qy	7312	ttcatgaacaagcttgaaaaaagttagagaagcgttcgaaggaagaagagcgtgacag	7371
Db	7889	TCCATGAGCAAGCTCGAGAAAGGTAAGAAACAGTCACAAAGAAAGAGCGCTGGCCAA	7948
Qy	7372	gggtggttbaagatggttcaacaggtctcccttgatataagcaccctcgttctgtctgt	7431
Db	7949	GGATGTTTGAAGGATG	



Db 1133 CACCCCGGAGGAGACAGAGATGCGGACATTACCGCTGGCCACAGTACGGCCCTCCAC 1192  
Qy 1196 gcaagaggagcaatctgcagccctccagatattgccccttctctcagatctctataa 1255  
Db 1193 ACCGGGGGGGCAATTGCACCCCTCCAGTATTGGCCCTTTCTTGAGATCTCTATTA 1252  
Qy 1256 ttgaaataacatccctctctcgaagatcccaagccccaaggggttggtga 1315  
Db 1253 TTGGAATACTATACATCCCTCTTTCTCGAGGATCCCAACGCCCTCACGGGGTGTGGA 1312  
Qy 1316 gtccctatgtctctcagcagcctacttggatgtatgtcaacagctcgtcagacact 1375  
Db 1313 GTCCCTTATGTTCTCTCACAGCCCTACTTGGATGATTGTCAACAGCTGCTCGAGACACT 1372  
Qy 1376 ctccaacccgaaggagagagagatctctttagaggtcagaataaagtctctgagc 1435  
Db 1373 CTTCAACACCGAGAGGAGAGAGAAATCTATTAGAGGCTTGAAATAATGTTCCCTGGGGC 1432  
Qy 1436 cgaagggcgacccaagcagcttgcaaatgagattgacatggagattcccttgactgccc 1495  
Db 1433 CGACGGGGGACCCACGGGGTTCGCAAAATGAGATTGACATGGGATTTCCCTTAACCTGCC 1492  
Qy 1496 cgttggagctacaacacgctgaagbaagagagcttgaaatctatcgcagagctc 1555  
Db 1493 CGGTTGGGACTACAAACGGGCTGAAGGTAGGAGAGCTTGAAATTCATGCGCCAGGCTCT 1552  
Qy 1556 ggtggcggtctcggggcgctcgaagaagcccaactaatttgctagaagtaagaaggt 1615  
Db 1553 GGTGGGGGCTTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTTAAGGTAGAGAGAT 1612  
Qy 1616 gatcagaggagacgaacgaactccctcgatattctctgagaggtctatgaaagccttcag 1675  
Db 1613 GATCAGAGGACGGAATGACCCCCCTCTGTTTCTTGAGAGGCTCTTGGAAGCCTTCAG 1672  
Qy 1676 gcggttaaccccttctgatacctactcagaagagcccaagagcctcagtgagccttc 1735  
Db 1673 GCGGTACACCCCTTTTATCCACCTCAGAGGCCCAAAAGCCCTCACTGCTTGCGCTT 1732  
Qy 1736 catgggcaatgcgctctgatactcagaagaactcagaagctcgaaggttacaaga 1795  
Db 1733 TATAGACACTCAGCCTTGTGATTTAGAAAGAACTTCACAGACTGGAAAGGCTTACAGGA 1792  
Qy 1796 ggcctgagttacgtatctagtgagagagagagaggtgatatcagaagggagacaga 1855  
Db 1793 GCGTGAAGTACTGATCTACTAGAGGAGGACGAGAAAGATATATTACAAAGAGACAGA 1852  
Qy 1856 aagggagaaagacagagaaagaaagagagagaaagggagaaagacgtgtag 1915  
Db 1853 AGAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAAAGGAGAAAGACGTAAATA 1912  
Qy 1916 acggcaagagaatcttgactaagatcttgccgagtggttgaaaggaagacagcaag 1975  
Db 1913 ACGGCAAGAGAAATTTGACTAGATCTTGCTGCTGCAAGTGTGAAGGAAAGCAATAC 1972  
Qy 1976 ggaagagagagagatcttagaanaatlagtcaagcccttagacagttcagaagactgg 2035  
Db 1973 GGAAGAGAGAGAGATTAGCAAAATTAGCTAGGCCCTTAACAGATCTCGGGAACCTGGG 2032  
Qy 2036 catatggacccaactcgcagaagcagtgctgctatgttaaagaaaagacacatgggc 2095  
Db 2033 CATATGAGACCCACTCGACAGAGACCAATGTCAATTGTAAAGAAAGGACACTGGGGC 2092  
Qy 2096 aaggaactgccccaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 2155  
Db 2093 AAGGAATCTCCCAAGAAAGGAAACAAAGAGCAAGGATCTTAAGCTTAAGAAAGATTA 2152  
Qy 2156 agatgaaggaagagaggtctcgagccctcccgagcccgaggtgaactttgaaggttga 2215  
Db 2153 AGATTAGGGAGACGGGGTTCGAGCCCTCCCTCCAGCCAGGAGATCTTTGAGAGGTGA 2212  
Qy 2216 ggggaacacagttgattccctggttgataccggagcggagcatcagtgctctacaacc 2275  
Db 2213 GGGCAACCAAGTGTCTCTGTTGATACCGAGCAAAATTAAGTCTACTACAGCC 2272  
Qy 2276 attagaataaagaaaaaatctcgttgatgagtgagtgagtgagtgagtgagtgagtg 2335  
Db 2273 ATTAGGAATCTAAATAATTAATAATCTGGGATGGGATGGGCAAGGGCAACAAAGTA 2332  
Qy 2336 tccatgactaccggaagacgcttgacttggaggttgaggggttaaccacacgttctc 2395  
Db 2333 TCCATGAGACTACCGAAGAACAGTGCATTGGGAGTGGGAGGGGTAACCACTGCTTCT 2392  
Qy 2396 gttatccctcagtgccagatccctctcaggttagagacttgcagaagatggagc 2455  
Db 2393 GGTATACCTGAGTGCCACACCCCTTTAGGTAGAGACTTATGACCAAGATGGAGAC 2452  
Qy 2456 lcaaatctcttgaaacaaagaaagaaaggtgtctgaaataaataaaccatcctgt 2515  
Db 2453 ACAATTTCTTTGAAACAAAGGAAACCAAGATGTCTGCAAAATACAAACCTATCTCT 2512  
Qy 2516 gttgacccctcaattagatgataatcagactatctccccaagtaagcctgatca 2575  
Db 2513 GTTGACCTCTCAATAGATGACGAATATGAGATATCTCCCTAGTAAACCTGATCA 2572  
Qy 2576 agatataagctcctggttgagagagttccccaagcctgagcaagaaccgagagatgg 2635  
Db 2573 AAATATACAAATCTGTTGGAAACAGTTTCCCAAGCCTGGCAGAAACCGCAGGATGG 2632  
Qy 2636 ttggcaaaagcaagttccccaaggttataactgaagggcagtgctacacagatc 2695  
Db 2633 TTTGGCAAAAGCAAGTTCCTCCCAAGATTTTCACTGAAGGCGCAGGACCAAGTGTG 2692  
Qy 2696 agtcaagacgtaccctctgagtagagaggtcgaagaagaaatttgccgcatgtcaag 2755  
Db 2693 AGTCACACAGTACCCCTTGAAGTAAAGAACCTCAAGAGGAATTCGGCGCATGTCCAAG 2752  
Qy 2756 attaatccaaagggcatctagttccctgctcgaatcccttggaaatacctccctgac 2815  
Db 2753 ATTTATCCACAGGGCATCTGATCTGCTGCTCCAAATCTCCCTGGAATCTCCCTGCTAC 2812  
Qy 2816 ggttgaagaccttggaacaaatgatalcagacagtaacagactggaagaggttcaataa 2875  
Db 2813 GGTGAAGAAAGCTGGGAGTAATGACTATCGACATACAGACTTGAGAGAGGTCAATAA 2872  
Qy 2876 aaggttgcaagacatacaccacaaggtcccgaaacctataactcttgagcctcc 2935  
Db 2873 ACGGTTGAGAGATATACACCAACAGTCCCAACCTTATTAACCTTGTGTGCTGCC 2932  
Qy 2936 gctgaagagactgtacacagatattgacttaagaatgctctctcgtcagatc 2995  
Db 2933 ACCCAAGGAGAGCTGTATACAGTATGAGCTTAAAGATGCCCTTCTCTGAGATT 2992  
Qy 2996 aacccccaatagccaacacttttgcctcgaatgagagatacagatccgggaagac 3055  
Db 2993 ACACCCCACTAGCAACACACTTTTTCCTTGCAATGAGAGATCCAGATACGGGAAGAC 3052  
Qy 3056 cgggacgtcactcgtggaactcgtcccaaggttcaagaactcccgacacatcttga 3115  
Db 3053 CGGGAGCTCACTGAGACCCGACACTGCCCAAGGTTCAAGAACTCCCGACATCTTTGA 3112  
Qy 3116 cgaagccctaaacagagagactggccaactcagaatcccaacacctcaagtgacccctc 3175  
Db 3113 CGAAGCCCTTACAGAGAACTGGGCAACTTCAGATATCAACACCTCAGTGAACCTGCT 3172  
Qy 3176 ccaatagtgatlaactgtctcgtggggagagcaacaaacagagctgcttaaggttac 3235  
Db 3173 CCACTAGTGTGATGACTGCTGCTGCGGGAGACCCACCAACAGAGACTTAAAGGCAC 3232  
Qy 3236 gaaggaactcctgctggaattgtcgaactaagctacagacccctcgtcgaagagccca 3295  
Db 3233 GAAGGCACTACTGCTGAATTTGCTGACCTAAGGCTACAGAGCTCTGCTAAGAGGCCA 3292  
Qy 3296 gatttgaagagagagtaacatacttgggtgacatttggcgggagcggagcagatgct 3355  
Db 3293 GATTTCAGAGAGAGATACATCTTGGGTACAGTTTGGCGGAGGGCAGGATGGCT 3352



QY 3356 gacggagacacgagaagaactygtatccagataccgcccccaacacagcaacagat 3415  
 DB 3353 GACGAGAGCACGGAAGAAACCTGTATGTCAGATACCGGCCCAACCCACACCAAAAT 3412  
 QY 3416 gagagagtttttggggaacgtgtatttgcacactgtgtatcccgaggttgcacct 3475  
 DB 3413 GAGAGAGTTTGGGGAGAGCTGTATTTGACAGCTGTGATCCCGGGGTTTGGACCTT 3472  
 QY 3476 agcagcccaacttaccgcttaacccaagaagaaggagatctctctgggtccctgaaga 3535  
 DB 3473 AGCAGCCCACTTACCGCTTACCGCTTATGTCAGATACCGGCCCAACCCACCAAAAT 3532  
 QY 3536 ccagaaggaatttgaatgatacaaaaaggccctgtgaagcagctgtcttggccctcc 3595  
 DB 3533 CCAGAGGACATTTGATGCTATCAAAAGGCGCTCTGAGCGCACCTCTCTGGCCCTCC 3592  
 QY 3596 tgaagtaactaacccttacccttattgtgatgaagtaaggaagtaagccagagat 3655  
 DB 3593 TGACCTAACTTAAACCTTTACCTTTATGTGATGACGCTAAGGAGATAGCCCGGGAGT 3652  
 QY 3656 tttaaccacaaccccttagagacatggaagagacctgttgcctacttcaaaagacttga 3715  
 DB 3653 TTTAACCCAAACCTTAGAGACCATGAGAAAGACTGTGCGCTACCTGTCAAAAGAGCTCGA 3712  
 QY 3716 tccgtgaacagtggttggcccgatagtctgaagactatgcagctgttggcctatgtgt 3775  
 DB 3713 TCCTGTACCCAGTGGTGGCCATATGCTTAAAGGCTATCCAGCTGTGGCCATATGTGT 3772  
 QY 3776 caagaagcctgaacaatgtacttttggacagataacttlaataagcccccaatgtat 3835  
 DB 3773 CAGAGACCTTGAACAATTTGACTTTGGACAGATATTAATTAATAGCCCCCATGTGAT 3832  
 QY 3836 ggaagaactcgttgcgcagcccccaagacagatgtatgaccaaagcccgatgaacacata 3895  
 DB 3833 GGAGAAACATCGTTCCGAGCCGCCAGACGATGATGACCAACGCCCCCATGACCCACTA 3892  
 QY 3896 tcaaaagctgtcttcaagagagaggttacttgcctcaccagccgttccaaacccgtc 3955  
 DB 3893 TCAAAAGCTGCTTCTCAAGAGAGGGTCACTGCTCCACACGCCGCTTCAAAACCTTGC 3952  
 QY 3956 cactctctgcctgaagactgtatgaacacagtgatcatatgtatgcaactattgtat 4015  
 DB 3953 CACTCTTCTGCTTGAAGAGACTGATGATGATGATGATGATGATGATGATGATGAT 4012  
 QY 4016 tgaagagactgggtgtccgcaagacacttaacagataccgctgtatgtgaagagtgctaac 4075  
 DB 4013 TGAGGAGACTGGGTCGCCGAAAGACCTTACAGACATACCGCTGATGAGAAAGTGTAAAC 4072  
 QY 4076 ctgttttcaactgaacgaagactatgtgtgtgaagtaagtaagatgtgtgtgtgtgtgt 4135  
 DB 4073 CTGCTTCACTGACGAGAGAGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4132  
 QY 4136 ggttgaacggagaccgcaacgacttgggcaagcagccttgcggaaggaacttaagcaaaa 4195  
 DB 4133 GGTGAGCGGAGACCCGACGATCTGGGCGCAGCCTGCGGAGAAAGACTTTCAGCAAAA 4192  
 QY 4196 ggtctgagctcatgtgcccctacgaagcttgcggtgtgcggaaggaatccataaact 4255  
 DB 4193 GGCTGAGCTCATGTGGCCCTACGCAAGCTTTGGGCTGGCCGGAAGGAAATTCATTAACAT 4252  
 QY 4256 ttataagacagcagtgatgacttltgacacttgcacagctacacggtgacatcatalaaca 4315  
 DB 4253 TTATAGGAGAGAGAGTATGCTTTGCGACTGACACGATACATGGGGCCATCTATTAACA 4312  
 QY 4316 aaggggtgtgcttaactcaagcagggaggaataaagaacaaagaggaattcttaagct 4375  
 DB 4313 AAGGGGTTGCTTACTCTGACGAGGGAGGAAATTAAGAACTTAAGGAAATTTCTAAGCTT 4372  
 QY 4376 attagaagcttactatttccaaaagagctagctatataactgtccctgtgaacagaa 4435  
 DB 4373 ATTAGAAGCCGTACATTTTACCAAAAAGGCTAGCTATTATACACTGTCTGTGACATCAGAA 4432

QY 4436 agccaaagatctcatatctagaggaagacagatgtgtgacggtgttccaaagcagcagc 4495  
 DB 4433 AGCTAAAGATTCATATTCAGAGAGAAACCAAGATGGGTGACCGGGTTTGCACAGAGGACG 4492  
 QY 4496 ccagctgttaacctcttgcctataatagaacgcccgaagcccaagacccagagcaga 4555  
 DB 4493 CCAGGGTGTAACTTCTGCTCTATTAATGAGAAATGCCCCAAAGGCCCAAGACGAGACA 4552  
 QY 4556 gtacacccctagaagacttgcgaagaagaataaagaatagaacagcttctgtgaactccga 4615  
 DB 4553 GTACACCCCTAGAAGACTGGCAAGATTAAGAAAGATAGACCACTTCTGACAGCTCCGGA 4612  
 QY 4616 gggagacctgtatcactcatatggaaggaagaatcttgcaccacaagaagaggttgaata 4675  
 DB 4613 AGGAGACTTGCTATACCTCAAGATGGGAAGAAATCTGCCCAAGAAAGAGGTTAGAATA 4672  
 QY 4676 tgtccaaagatacatcgttcaacccaactagaactaaacaccttcaagcagttgttgcg 4735  
 DB 4673 TGTCCAAACATATCATCTGTCTAACCCACCTAGGAACTTAAACACTCTCACAGTTGGTCCAG 4732  
 QY 4736 aacatcccttcatcagcttgccttgaagcagagagtggtgtgactgtgtgttcaacatgt 4795  
 DB 4733 AACATCCCTTATCATCTTGTGAGCTACAGAGAGTGGCTGACTGGTGGTCAACATTTG 4792  
 QY 4796 tgtcccttgcagctgttgaatgttaactccttccagaatacctccagaagaagactaag 4855  
 DB 4793 TGTGCCCTTGCCAGCTGGTTAATGCTAATCTTCCAAATGCTCCAGAGGAAAGACTAAG 4852  
 QY 4856 gggagacccacccagcgtctaccttgcggaagtgagcttcaaggttaagccggttaata 4915  
 DB 4853 GGGAAAGCCACCCAGGCGCTCTACTGGGAAGTGGACTTCACTGAGATTAAGCCGCTTAATA 4912  
 QY 4916 cggaaacaatatctattgttlltltgaacaccttllcaaga ttggttgaagcttacc 4975  
 DB 4913 CGGAACAATAATFACTATTGGTTTGTGAGACACTTTTTCAGATGGGTAGAGCTTATCC 4972  
 QY 4976 tactaagaagagactcaacgltgtgtgclaaagaaatactcgtgaggaatlltccag 5035  
 DB 4973 TACTAAGAAGAGACTTCAACCGTGGTGTAAATAAAATCTGGAAGAAATTTTTCAG 5032  
 QY 5036 atttggaatcctaagtaataggttcagagaatgtgtcgaagcttctgttccaggtlaag 5095  
 DB 5033 ATTTGGAATACCTTAAGGTAAATAGGGTCAAGATGTGCAAGCTTTTGTTCCTCAAGGTAA 5092  
 QY 5096 tcaaggaactgcacaagatatltggtggtatgtatggaacttgcattgtgtcatacagacc 5155  
 DB 5093 TCAGGAGACTGGCCCAAGATATTGGGATTTGATGGAACCTGCATGTGCTATACAGACCCCA 5152  
 QY 5156 aagctcagagcaggttagagagatgaatgaacacataaagagaccccttaccgaattgac 5215  
 DB 5153 AAGCTCAGGACAGGTAGAGAGATGAATGAAACCATTAAGAGACCCCTTACTTAATTTGAC 5212  
 QY 5216 cacagagactgcactaatgaattgattgtgtctctccttgccttgccttlltagagttga 5275  
 DB 5213 CGCGAGACTGGCTTAAATGATTTGATAGCTCTCTGCTTGTGTGTGTGTGTGTGTGTGT 5272  
 QY 5276 gaacaccccttgaacagtttgggtgtgacccctatgaattgtctctaaggggagccccccc 5335  
 DB 5273 GAACACCCCTGACAGATTTGGGCTGACCCCTATGAATTACTTACACGGGGACCCGCCCC 5332  
 QY 5336 gttggcagaatltgccttgcataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5395  
 DB 5333 ATTGGTAAATAATTTGCTTGTGACATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5392  
 QY 5396 taagctcaagcgtcaggt 5455  
 DB 5393 TAGGCTCAAGGACACTTGTGATGGGTGAGCAACGAGCGTGGGACACTCCGGAGGCGCTA 5452  
 QY 5456 etc---aggagagacttgcaggttccacatcgtcttccaaagtgtgagattcagctatgt 5512  
 DB 5453 CTCAGAGAGAGAGACTTGCAGATCCCAATCGTTTCCAAAGTGGAGATTCAGCTACGT 5512  
 QY 5513 tagagccacgltgcaggaacccctgcagactcgtltggaaggaagacttattcgtactt 5572





OY	7712	ttaacagaccctgctgcgcagtaaaaggtagaaggtcacactlccatctgttccagg	7771
Db	7658	TAGGTAAAGANTTAGGTTTTTTTGGCTGTTTAAAATATCGSGAAGTAATAATAGCCCTGAG	7717
OY	7772	gccctgatccctggcct-aagtaaatacacagaatatgaattgac---taalcgcttat	7826
Db	7718	TACATGTCCTCTAGGCATGAACCTCCTTGAAACTATTGAGATTAACAAGAATAAGGAGTTT	7777
OY	7827	ctggaattctgtaaacatctaactgycacatalagaagaattgatlacacattgacgccttag	7886
Db	7778	CTAAAGCTGTTGTTTAGCTTCGTETAACCTGGTTCCGCCATTAAGATGTTTAAATGTTGAT	7837
OY	7887	tgaccatattcaacatcgacatcgttcacactctgccaggagcccacgacagtgcgacctcc	7946
Db	7838	ACACATATCTTGGTGACAAACATGTCCTCCACCACCCCGAACAATCGCAAAATGTGTAACTC	7897
OY	7947	ggaagcatlttaaatgattggtccacagagcgcggcctcgtatatttaaaatgattg	8006
Db	7898	TAAACCATTTTAATTAATTAATTTGGTCCACGAAGCGGGCGCTCGAAGTTTAAATTGACATG	7957
OY	8007	gtccaatgagcgcgcggcctcgtatatatttaaatgattggttttgagcagcagcgcttg	8066
Db	7958	GTTTGT-----GATATTTTGAATCATTTGGTTTAAAGCGGGGCTTG	8002
OY	8067	ttgtgacccccataaagaatgtcccgatcttcgcgaactcggagcgcgaatcctctaacctg	8126
Db	8003	TTGTGAACCCCATTAAGAAGTGTCGCCGACTCCACACTCGGGGCGCAGTCTCTAACCCCTG	8062
OY	8127	cgtggtgtacgaactgtgggcacccaagcgcgcttgtgaataaaaatccctctgctgttgcat	8186
Db	8063	CGGGGTGTACGACTGTGGGCCCCAGCGGCGCTTGAAATAATCCTCTGCTGTTTGCAAT	8122
OY	8187	caaaaaaaa	8196
Db	8123	CAAAAAAA	8132
RESULT	10		
AF038600			
LOCUS	AF038600	8132 bp	mRNA
DEFINITION	Sus scrofa porcine endogenous retrovirus PERV-MSL mRNA, complete sequence.		
ACCESSION	AF038600		
VERSION	AF038600.1	GI:3133301	
KEYWORDS			
SOURCE	pig.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.		
AUTHORS	1 (bases 1 to 8132) Akizoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.		
TITLE	Identification of a full-length cDNA for an endogenous retrovirus of miniature swine		
JOURNAL	J. Virol. 72 (5), 4503-4507 (1998)		
MEDLINE	98216827		
REFERENCE	2 (bases 1 to 8132) Fishman,J.A.		
AUTHORS	Direct Submission Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA		
TITLE	3 (bases 1 to 8132) Fishman,J.A.		
JOURNAL	Direct Submission Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA		
REMARK	Sequence update by submitter On May 16, 1998 this sequence version replaced gi:3116441. Location/Qualifiers		
COMMENT	1. 8132		
FEATURES	/organism="Sus scrofa" /strain="miniature swine"		
SOURCE			

```

/db xref="taxon:9823"
/cell-type="lymphocytes"
/note="porcine endogenous retrovirus PERV-MSL"
1..564
585..2159
/codon_start=1
/product="gag protein"
/protein_id="AAC16763.1"
/db_xref="GI:3116442"
/translation="MGQVYTPPLSLTLDHMTVEKSRANHLSVQYKKGMPQTECVSEMP
KRPDGPSPRIALIGEKKKHSAEKVPSPHLYPFLPEEPAPREPOSVPPYLAOGAARG
KRVNPGDAPRPRLOINEIMGPEPLRPGMDYNATRESLKLYRQALVGLGRASRR
KTIQRLKEGOEAELRDLYAEAEVYVKRTPEEREDRKRREERERENRKNKLT
KILAAVECKSNTEREDFRKII RSGPRGSGNLGNRTPLDKQCAVCYKEGHWARNCPK
KNGKGRILLAEEDKD"
2307..5744
/codon_start=1
/product="pol protein"
/protein_id="AAC16764.1"
/db_xref="GI:3133302"
/translation="MGATGQOQVPMWTRRTYDVGVRWTHSVLPTECPAFLGRDL
TKMGAISEBOGKEVVSANKRPITVLTQLDDEYRLSPHYKPDQNIQWLBPPQAP
AETAGAGLAKOVPPOVYIOLKASATPVSRQYPLSKDAEGIRPHVRLIQOGLIVPQ
SPNPTLLPVRKPGTNDYRPVDLREVRKVDIHPTVLPVNNLCALPDRSMYTVL
DLKDAFECLDLHPTSQPLFAFEWRDGTGQGLTWTLRPOGKSNPTFDEMLHRLD
ANRDIQHOVTLIOYDLDLLAGATKQDCEGRKALLLESDYGVASAKQAICRE
VYLVGLSLRDGQWMLTEBAKTKYVQIPAPYPAKOMEPELGTAGCFYMLIPGATLAP
LYPLTEKEGEFSAPEHQKAFDAIKKALLSAPALALPDYTKPTLYLDEKKGARVYL
TQTLGWRKRPVAYLSKRLDPVASGMPICLKALAAVA1LVKDAKLTGQNTVIAHPA
LENIIVGRPPRMMTNARKMTYHOSLILTERVTFAPPALNATLILPETEPPVNDHQ
LILKEEGVRKDLIDPLTGEVLTFWTDGSSVYVEGRMAGAIVADGRTIMASRELG
TSQKREKMLTQALRIAEKGSINITYDSYATATVGHGAIYKQGLITLSAGREIKN
KEETLSILEAVHLPKRLATIHCPGHOKAKLTISRGMDADRVAKQAAQVNLPITEM
PKAPEREPRRTLEDWQETIKIDQSETPBEQTCYTSQKELLPKKEGLEVQNIHLRTH
LGRKILQDLRTSDPYHVLRLPGVADSVKHCVCQULNANPSSMPPEKRLRGSHPGH
MEWDFTEFPAKYGKNYLILFVVDTFSGMVAPYKPKETSVVAK1LLETLPFRGLPK
VISEDNGEPAVHAWSQIAKILG1DMKLCIYAPSSQGOVERNKK1LETLPTLTAET
LGNADMLALPEVLFVRNTPGOGELPTELYLCCGPRPLYIAVSHVADLLSQPLSR
LKALENVRQAMKQLREAVSYSGGDLQIPHEFGYDGSVYRHHAGNLETWKPKYVL
LTPPLAVKVEGISTWTHASHVYKPAAPPDSQWKAKEITNPLKLLHRVYVPSVNNLSQ"
5620..7336
/codon_start=1
/product="env protein"
/protein_id="AAC16765.1"
/db_xref="GI:3133303"
/translation="MHPTLNRHLPIRGKRPKRKLTPLSFASIAWELTSLTSQNGM
RIGDLSNMRPLSLTWLITDSGTGIN1NNQGEPLSTAMPDLVLCRLSVTSLTSSP
DILHAHGFVPCGPPRNKSGHCGNDRDFCKOENNVTSNDYMWIPSSQDRAVSYVGN
TYTSSGFNFYLTWIRNNGSPKSCSPDLDYLIISFENKCEKNTLKWVNGMWSYVY
SGKQPSILITRKLINQLEPPMAIGPMTVLTGQPPGQSGPSNLTSGDPTSSST
TKMGAKFLILQCAFQALNSTTPGASPCMLCLASGPYPEGARKEPFWYKTHRQD
CTWMSOKSLTUTEASGKGTICAKYPSHONCHNTEAPFNQTSQYTVPYDMMACN
TGLPPLCVSTLVNQTKDFCIMVQIVPVVYVPEKAL1LDEYDRNHRKRREPLSLTVA
MLGEGVAGAVGTGAALVGTGPOULETSLNHRVTDGLQALEKSYSNLESLTSLSE
VVLQNRGLDLFLKEGGLCVALKCECCCFYVDHSGA1RDSNNK1RELELRKRREKETT
OGMEGKFNRSIMATLTLTSLATGLPLVLLLLTVGPCIINKLIAFIERISAVQIMVIL
RQYQSPSSRRACR"
7518..7976

```

QY 61 aagaccgttcgttgatgatt-tgggtgtcgcctcttcagaccggagaggga 119  
 |||||  
 Db 63 AAGACGGCTTCTGTGATGATTAAGGGAGTGCCTTTTCGAGGCTGGAGG----- 115  
 QY 120 tttgtcttctacgtcttcatttggtgtcgttgccggaaatccttgaccacccct 179  
 |||||  
 Db 116 --TTCTTTTGGTGGCTTACATTTGGGGGCTGTCGGGATCTGTCCGGCCACCCCTA 173  
 QY 180 aacaccgagaaacagcttggagta-aaggaatccctcttgaaagctgtg--tggtc 236  
 |||||  
 Db 174 ACACCCGAGAACGACTTGGAGTAAAGATCCCTTTTAAGGTGATGATGATAC 233  
 QY 237 ggcggcgctctctgtcttgatgtctgttcttcgtgtgagcgcttcogtttgacgtg 296  
 |||||  
 Db 234 GGGCGGCGTCTGTTCTGATGATGCTGTTTTCAGTGTGTCGGCTTGGCTTTCAGCGTG 293  
 QY 297 tccctcagaacgttgaagacttgatcagacagagctgtctggaggatcaag 356  
 |||||  
 Db 294 TCCTTCAGGCGCTGAAGGCTGGGGACTGTGATACAGACGATGCTTAGAGGATCACAG 353  
 QY 357 gctgcaccccttgaggagagcccgaggagtgaggagacagagcgctgtgtgtc 416  
 |||||  
 Db 354 GCTGCTGCTTGGGGAGCGCCCGGAGAGTGAGAGACCAAGGACGCTGTGTCTTC 413  
 QY 417 tactgtcgtcagagagagagcttctgtgttgaaagcgaaagctccctccgagcg 476  
 |||||  
 Db 414 TACTGTGCTGAGAGGACCGAATTTGTTGCTGAAGCGAAGCTTCCCTCCCGACCG 473  
 QY 477 tccgaactctttgctgtcttgaaagacgagcggtgtgtgtctgtctgtgtgt 536  
 |||||  
 Db 474 TCCGACTCTTGTGCTGTGTGGAAGAGCGTGAAGCGGTGACGTGTGTGTGTGTGT 533  
 QY 537 gttctgtctgtgtgtcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596  
 |||||  
 Db 534 GTTCTGT 593  
 QY 597 aacgtgaacacccctctgactgtgactgtgacatgttgatagatccagagct 656  
 |||||  
 Db 594 ACGGTGAGGACCCCTCTTATGTTGATCTGACACATTTGAGTGAATTAATCCAGGGCT 653  
 QY 657 cataattgtcagttcaggttgaagaagagacttgacagcttctgtcctgtgtgt 716  
 |||||  
 Db 654 CATATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713  
 QY 717 ccaacattcgaatgttgatgagccctcagagagagaccttaattctgaatlatccgt 776  
 |||||  
 Db 714 CCGACATTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 773  
 QY 777 gtaaggaatcattcttcagactgagcccgctcctcactcactcactcactcactc 836  
 |||||  
 Db 774 GTTAAGCAGTATTTTTCAGACGAGACCGGCTCTCTATCCGATCCAGGAGCCCTATATC 833  
 QY 837 cttaactgtgcagaagatttgcagaagatccctccgacatggttlaaacatgtctaaataa 896  
 |||||  
 Db 834 CTTAACGTGAGAAATTTGGCAGAGATCTCCGCAATGGGTAAACCATGCTGATTAAG 893  
 QY 897 ccaagaagacagatcccgaaatccgtgtcttgaggagagaaacaaacactcgcgaa 956  
 |||||  
 Db 894 CCAAGAAAGCAGGTCCCGCAATTTCTGGCTTTGGAGAGAAACAAACCTCGGCTGAA 953  
 QY 957 aaagtgcagccctctcctcgtactaaccagagatcgagagcgccagcttgccgga 1016  
 |||||  
 Db 954 AAAGTCAAGCCCTC-TCCTCATATCTAACCCGAGATTGAGAGACCAACGGCTTGGCCGA 1012  
 QY 1017 accccaactgttcccccaacccttactcagacagaggtgtctgtgtgtgtgtgtgtgt 1075  
 |||||  
 Db 1013 ACCCCATCTGTTCCGCCACCCCTTATCTGCGACAGGGTGGCGGAGGAGACCTTTTGC 1072  
 QY 1076 cccctcctgagagctcaggtgtgagagagactgtccgagactcagagcgagagagagc 1135  
 |||||  
 Db 1073 CCTCTCTGAGCTCCGGCGGTGGAGGACCTGCTGAGGGAGCTCGGAGCCGAGGGGCGC 1132

QY 1136 caaccggagcggagacagagatcgacataaacgcgtgcacactatgcccctccat 1195  
 |||||  
 Db 1133 CACCCGGAGGGAGACAGACGATTCGACATTTACCGTGTGGCAGCTAGACGCTCCAC 1192  
 QY 1196 gccagggggcaattgcaagccctccagatattggcccttctctctcagatcctctaa 1255  
 |||||  
 Db 1193 ACCGGGGGCGCAATTTGACGCCCTTCAGATTTGGCCCTTTCTCTGACATGCTATTA 1252  
 QY 1256 ttggaatactaacatccctctctcggagagatcccccaagcctcaggggtgtgtga 1315  
 |||||  
 Db 1253 TTGGAAACTTAACATCCCTCTTCTGGAGATCCCAACGCCCTCAGCGGGTGTGGA 1312  
 QY 1316 gtccctatgtctcctcaccagcctacttgatgtatgtgtgtgtgtgtgtgtgtgtgt 1375  
 |||||  
 Db 1313 GTCCCTTATGTTCTCTACACAGCCTACTTGGAGATGTTGCAACAGTGTGTGAGACACT 1372  
 QY 1376 cttaacaaacgagagagagagaaatctcgttagagagctagaagaaatgtctctgggc 1435  
 |||||  
 Db 1373 CTTCAACAACGAGAGAGAGAGAAATTTATTAAGAGCTAAGAAATAATGTTCTGGGCC 1432  
 QY 1436 cgaagcgagacccagcaggtgtgcaaaatgagatgtgaatgtgtgtgtgtgtgtgtgt 1495  
 |||||  
 Db 1433 CGAGGGCGACCAACGCGGTTTGCAAAATGAGATTGACATGAGATTTCCTTAACGCGCC 1492  
 QY 1496 cgttggagactaacaacagcgtgaagtgagagagcttgaaatctatccagagctct 1555  
 |||||  
 Db 1493 CGGTTGGAGCTACAAACGCGCTGAAGGTAGAGAGACTTGAATAATATATCCAGGCTCT 1552  
 QY 1556 ggtggcggtctccggggcgctcgaagagcgcccaatattgtgtcgaagtaagagag 1615  
 |||||  
 Db 1553 GGTGGCGGCTCTCGGGGCCCTCAAGACGCCCTCAATTTGGCTAAGTAAGAGAGAT 1612  
 QY 1616 gatgcagggagcgaacgaacccctccglatcttcttgaggagctatagaagcctcag 1675  
 |||||  
 Db 1613 GATGACGAGGACCGAATGAACCCCTCTGTTTCTTGAGAGCTCTTGAAAGCTTCAG 1672  
 QY 1676 gcggttcaacccctttgactcactcagagagcccaagagcctcagtgccctgcct 1735  
 |||||  
 Db 1673 GCGGTACACCCCTTTGATCCCACTCAGAGGCGCCAAAAGGCTCAGTGTGCTTGGCTT 1732  
 QY 1736 catgggagctggt 1795  
 |||||  
 Db 1733 TATAGACAGTCACTGCTTGATATTAAGAAAGACTTCAGACACTGGAAGGCTTACAGA 1792  
 QY 1796 ggtcaggttcaagctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1855  
 |||||  
 Db 1793 GGTGAGTGTACGTGATCTGTGAGAGAGCAGAGAAATGATTTTCAAAAGGAGAGAGA 1852  
 QY 1856 ag 1915  
 |||||  
 Db 1853 AGAAGAAAG 1912  
 QY 1916 acggcagaag 1975  
 |||||  
 Db 1913 ACGGCAAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1972  
 QY 1976 gtag 2035  
 |||||  
 Db 1973 GGAAGAGAGAGAGATTTAGGAAATTAAGTCAAGGCGCTTGAAGAGAGAGAGAGAG 2032  
 QY 2036 caatagagacccacacacagagagagagagagagagagagagagagagagagagag 2095  
 |||||  
 Db 2033 CAATAGGAGACCCACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092  
 QY 2096 aaggaactgcccccaag 2155  
 |||||  
 Db 2093 AAGGAAGTGGCCCAAG 2152  
 QY 2156 agatag 2215  
 |||||  
 Db 2153 AGATTAGGAG 2212  
 QY 2216 gggggaaccagttgagttcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2275

|||||  
Db 2213 GGGGACACCACTTGAATTCCTGGTTGATACCGGACGAAACATTCATGCTACTACAGCC 2272  
Qy 2276 attagaaacataaagaaaaaaalocctgggtgagtggtgcaagggacggcaagta 2335  
|||||  
Db 2273 ATTGAGAAACCTAAAGATATAAAATCTGGGTGATGGTCCACAGGGCAACAAACGTA 2332  
Qy 2336 tccatgagtaaccggaagaaacgttgcattgggagtggtggcggttaaccatcgttct 2395  
|||||  
Db 2333 TCATAGGACTACCGAAACAGTGTGACTGGAGTGGAGGGGATACCCACTGTTTCT 2332  
Qy 2396 ggtacccctgagtgccagtaacccttctagtagagtaacttactgaacaaagtgagc 2455  
|||||  
Db 2393 GGTCTACTAGTGTCCAGACACCCCTTGTAGTAGAGACTTATTGACCAAGATGGGAGC 2452  
Qy 2456 tcaaatctcttggacaagaagacagaagtgctgtgtaatacaaaaccatcaatgt 2515  
|||||  
Db 2453 ACAATTTCTTTGMAACAAGGMAACGAGAAGTGTCTGCAATTAACAACCTATACCTGT 2512  
Qy 2516 gttgacctccaatagatgagatgacatgacatattctcccaagtaaaagcttgatca 2575  
|||||  
Db 2513 GTTGACCTCCATTAGATGAGATGAGATGACTACTCTCCCTAGTAAAGCCTGATCA 2572  
Qy 2576 agataacagtcctggttggagcaagttccccaagcctgggcagaacccgcaaggtatgg 2635  
|||||  
Db 2573 AAATATACATTTCTGTTGGAAACAGTTTCCCAAGCCTGGGACAGAACCCGAGGATGG 2632  
Qy 2636 ttggcaagcaagatgccccaagttatcaactgaagccagtgctacacagatc 2695  
|||||  
Db 2633 TTTGGCAAGCAAGTTCCTCCCAACAAGTATTCACCTGAAGGCGATGCAACAGATCTC 2692  
Qy 2696 agtcgaagcaagtaacccttgagtagagagcctgcagaagaagtcttgccgcaatgtaag 2755  
|||||  
Db 2693 ACTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGCAATGCCAAG 2752  
Qy 2756 attaaaccaagggacatcctagtcctgtcccaatcccttggaaactcctccctgtaacc 2815  
|||||  
Db 2753 ATTATTCACCAAGGGGATCTGATGTCCTGATCAATCTCCCTGAGATCTCCCTGCTACG 2812  
Qy 2816 ggttaggaagcctggagcaaatgatatccagtaacagtaagacttgaaagaggttcaataa 2875  
|||||  
Db 2813 GGTGAGAAAGCCTGGAGACTAATGACTATCGACAGTACAGCACTTGAAGAGGTCAATTA 2872  
Qy 2876 aaggtgtgcaagacatacaccacgaagtcgccgaacccttataacccttgaagcgccctcc 2935  
|||||  
Db 2873 ACGGAGTACAGATATACCCCAACAGTCCCGAACCTTATACCTCTTGTGTGCTCTCC 2932  
Qy 2936 gcttgaacggaactggtacacagtaattgactttaaagatgctctcttctgctgagat 2995  
|||||  
Db 2933 ACCCCACGAGAGCTGATATAGATTTAGATTTAAAGATGCTCTTCTGCTGAGATT 2992  
Qy 2996 acaaccactagcaacacacacttctgcttgaagtgagagatccagtgtaacgggaagac 3055  
|||||  
Db 2993 ACCCCACTAGCAACCACTTTTGGCTTCCAAATGGAAGATCCAGGATCGGGAAGAAC 3052  
Qy 3056 cggagcagctcacttgaccagcagctgcccgaaggttcaagaaactccccaacatcttga 3115  
|||||  
Db 3053 CGGGCAGCTCAGCTGAGACCCGACTGCCCCAAGGTTCAAGAACTCCCGACATCTTTGA 3112  
Qy 3116 cgaagcctcacaaggaaccttggccaacttcagatccacaacccctcaagtgtaaccctct 3175  
|||||  
Db 3113 CGAAGCCCTACACAGAGACTGCGCAACTTCAGGATCCAAACCTCAGGTGACCTCTCT 3172  
Qy 3176 ccagtaagtgatgagctcttctgcgaggaagccacaacagagctgcttgaaggtac 3235  
|||||  
Db 3173 CCAGTACGTGATGACTCTCTTGCGGAGAGCCACCAACAGAGATGCTTGAAGAGCAC 3232  
Qy 3236 gaaggaactactgctgaattgtctgaactaggtacagaagcctctgtctaaagaagccca 3295  
|||||  
Db 3233 GAAGGACATATCTGCTGGAATTTGCTGACTAGGCTACAGAGCTCTGCTTAAGAGGCCCA 3292  
Qy 3296 gatttgcaggaagagtaacataacttgggtgtaagttcggggagcgagcagatgct 3355  
|||||

---

Db 3293 GATTTCAGAGAGAGATTAACATTTCTGGGTAACAGTTTACGGGACGGGACGATGGCT 3352  
Qy 3356 gacagaggaacaggaagaacacttgatccagataccggcccaacacacagcaaacag 3415  
|||||  
Db 3353 GACGAGAGCACGGAAGAAACTGTACTCAGATACCGGGCCCAACACAGCAACCAAT 3412  
Qy 3416 gaagagtttlltggagcaagcttgatatttgacagcttgatccgggttltgagact 3475  
|||||  
Db 3413 GAGAGAGTTTGGGACAGACTGGATTTTGACAGACTGTGGAATCCCGGGGTTGGGACCTT 3472  
Qy 3476 agcaagcccaacttacccttaacaaagaagaagggatcttctgggtccctgaaga 3535  
|||||  
Db 3473 AGCAGCCCACTTACCCTGTAACAAAGAAAGGGAATTTCTCGGCTCTTAGCA 3532  
Qy 3536 ccagaaggaacttgatgacttcaaaaagccctgcctgagcgcaacctgtctgcccctcc 3595  
|||||  
Db 3533 CCAGAGGCACTTTGATGCTTCAAAAAGGCCCTGCTGAGGCGCACGCTGTGGCCCTCC 3592  
Qy 3596 tgaactactaaaccccttaacccttattgtgagtgaagcgtlaaggaagtagccgagagt 3655  
|||||  
Db 3593 TGACGTAACTAAACCTTTACCTTTATGTGATGAGCGTAAGGAGTACCCCGGAGAGT 3652  
Qy 3656 tttaacccaacccctggagccatggagaggaacctgtgtgcttactgttcaagaagctga 3715  
|||||  
Db 3653 TTTAACCCAAACCTTAGGACATGAGAGAACACCTGTGCTACCTGTCAANAGAGCTCGA 3712  
Qy 3716 tccgtgtagcagtgatgtgcccgtatgtctgaagagctatgcagctgtgtgcca tactgt 3775  
|||||  
Db 3713 TCCTGTAGCCAGTGTGTGGCCATATGCTCGAAGGCTATGCGAGCTGTGGCCATACGT 3772  
Qy 3776 caaggaagcttgacaacttgacttggagcaagaataactgtataagcccccaatgact 3835  
|||||  
Db 3773 CAAGGAGCTGACAAATTTGACTTTGGAGCAAGATATTAACGTATATACCCCTCATGATP 3832  
Qy 3836 ggaagaatcgttgcggaagcccccagaacagatgagtaacaaagcccgca tgaaccca 3895  
|||||  
Db 3833 GGAGAAATGCTTGGAGGCCCCCAAGCCATGATGATCAACAGCCCGCATGACCACTA 3892  
Qy 3896 tcaagcctgcttctcacaagagaggtcacttgcctcacaacagccgctcacaaccctgc 3955  
|||||  
Db 3893 TCAGAGCTGCTTCTCAGAGAGAGGTACAGTTCGCTCCACAGCCCTCTCAACCTGC 3952  
Qy 3956 cactcttctgcttgaagagacttgataccagtgacttcaatgatttgcataactatgat 4015  
|||||  
Db 3953 CACTCTTCTGCTCAAGAGACTGATGAACCAAGTACATGATTTGCCATCACTATTGAT 4012  
Qy 4016 tgaaggaagctgggttccgcaaggaaccttaagacataaccgtgaacttgagaagtgaac 4075  
|||||  
Db 4013 TGAGGAGACTGGGTCCGCAAGGACTTACAGACATACCGCTGACTGAGAAAGTCTAAC 4072  
Qy 4076 ctgttcaactgaaggaagcaactatgtgtgtaaggtlaagaga tggcctgggcygag 4135  
|||||  
Db 4073 CTGGTTCACTGACGGAAACACTTATGTGTCGAAGGTAAAGAGATGCTGGGCGGCGGT 4132  
Qy 4136 ggtggaaggaacccgcaagatcttggtccagcagccttgcggaaggaacttcagggcaaa 4195  
|||||  
Db 4133 GGTGAGACGGGACCCGACAGATCTGGGCACAGGCTCCCGGAGAACTTACGACAAA 4192  
Qy 4196 ggtcgaagctatggccctcacaagccttgcggtcggcctggccgaaggaatccataaacat 4255  
|||||  
Db 4193 GGTGAGCTATGGGCTTACGCAAGCAAGCTTTGCGGCTGGCCGAAGGAATTCATAACAT 4252  
Qy 4256 tlataaggaacagatgacttgccttgcagctgacacagtlacaggggacatctataaca 4315  
|||||  
Db 4253 TTATATCGACACGAGGTATGCTTTGCGACTGACAGGTATATGAGGCTCATCTTTAACA 4312  
Qy 4316 aaggggttltgttacttaagcaagggaggaataaagaacaagaaggaatcttaagcct 4375  
|||||  
Db 4313 AAGGGGTTGCTTACCTTACGAGGAGGGAATTAAGACAAAGGAATTTCTTAACCT 4372  
Qy 4376 attaaagccttaacttgcacaaaagcctagctatataacagcttcttgagata tcagaa 4435  
|||||  
Db 4373 ATTGAACGCTGATCTTACCAAAAAGGCTAGCTATTATATACACTGCTCGACATCGAA 4432

4436 agccaaagatctcatalctagagggagaccagatgctgacgggttcccaagcaagcagc 4495  
4433 AGCTAAAGATCTCATATTCAGAGAAACAGATGGGTGACGGGTTCACACAGACAGC 4492  
4496 ccaagctgttaacctctgctctaaatagaacgcccagaagcccagaacccagaacga 4555  
4493 CCAGGCTGTAACTCTTCCCTATATATGAATAATGCCCAAGGCCAGAACCCAGACACA 4552  
4556 gtacacccctagagacttgcagagataaagaatagaacggttctctgagactccga 4615  
4553 GTACACCCCTAGAGACTGGCAAGATATAAAGATAGACAGTCTCTGAGACTCCGGA 4612  
4616 gggagactgtctaacctatactggaaggaatccctgcccagaagaaggttagaata 4675  
4613 AGGAGACTCTGTACCTCAGATGGAGAAATCCTGCCCAACAAGAGGTTGATTA 4672  
4676 tctcaacagatactgctctacccacccactaggaactaaacactgcagactgtgtag 4735  
4673 TCTCAACAGATACATCGCTTAACCCACTAGAACTAAACACCTGCAGCAGTTGTCAG 4732  
4736 aacatccctctactgtctctgagctacccagagtgctgactgctgctcaaacatg 4795  
4733 AACATCCCTTATCATGTTCTAGGCTACAGAGATGGCTGACTGCTGCTCAACATG 4792  
4796 tgtccctgacagctgtgttaactgctacccactccagaatacctccagaagaagactaag 4855  
4793 TGTGCTGCTGCGAGCTGTTAATGCTAATCTCCAGAAATGCTCCAGGAGAGAGACTAG 4852  
4856 gggagagcaccacagcgtctactcaggaagtgagactcactgaagttaagccgctaaata 4915  
4853 GGGAGACCAACAGGCGCTCACTGAGAGTGGACTCTAGAGTAAGGCGGTTAATA 4912  
4916 cggaaacaatactatctgtgttctgtagacaccttccagaatggtgtagaggtctatcc 4975  
4913 CGGAACAATAATCTATTGTTTGTAGACACCTTTCAAGATGGGTAGAGGCTTATCC 4972  
4976 tactagaagaagacttaaccgtgtgtgctcaagaataactctgtagaagaatlttccaag 5035  
4973 TACTAGAAGAAGACTTCAACCGTGGTGAATAAATACTGGAAGAAATTTTCCAG 5032  
5036 attggaatacctaaggaatgaaggttcacagactgctcactgcttctgcccagagtaag 5095  
5033 ATTGGATACCTTAAGGTAATAGGTGACAGCAATGCTCCACTTTTGTGCGCCAGTAG 5092  
5096 tcaaggactgtgccaagatatgtgggattgtatgtgaactgcattgtgcatcacagaccga 5155  
5093 TCAGGACTGGCCAAAGATATTGGGATTGATTGAACCTGATTTGTGCATACAGACCCA 5152  
5156 aagctcagagacaggtagagagatgaaatagaacacattaaagaagacccctaccaattgac 5215  
5153 AAGCTCAGGAGAGGTAGAGAGATGAATAGAACATTAAAGAGACCTTACTAATATGAC 5212  
5216 caccagagactgtcataatgtatgtatgctcctcctgcttctgttctttagggtag 5275  
5213 CGCGAGAGCTGGCTTATATGATGTGATAGCTCTCTCTCTCTTGTCTTTTATAGGGTAG 5272  
5276 gaacacccctgtgacagcttgggtgtgacccctatgaattgtctctacagggagacccccc 5335  
5273 GAACACCCCTGTGACAGTTTGGGCTGACCCCTATGATTACTCTACGGGGAGACCCCCC 5332  
5336 gtgagcagaatgtgcttgcataatagtgctgtatgtctgtcttccacagccttgtctc 5395  
5333 ATTGTGAAATGCTCTTCTATACATAGTGTGATGTCTCTTCCAGGCTTTGTGTC 5392  
5396 taagctcaagcgctcagctgtgtgagcgagcgctggaagcagctccggagagccata 5455  
5393 TAGGCTCAAGGCACTTAGTGGGTGAGACACAGCGCTGAGGCAACTCCGGAGGCGCTA 5452  
5456 ctc---aggaagagactgtcaagttccacactgctcccaagttgagatcagctatgt 5512  
5453 CTCAGGAGGAGAGACTTGCAGATCCACATCGTTTCCAAAGTGGAGATTCACTTACGT 5512  
5513 tagagccacagctgcaggaacactcagagactcgtgtgaaggaacttactctgtacttt 5572  
5513 TAGAGCCACCGTGCAGAGAAACCTGAGACTGGGTGAGAGGCGCTTATGCTGATCTTT 5572  
5573 gaccacacacagctgtgaaagtcgaaggaatccccacactgtgatactatccacgt 5632  
5573 GACCACACCAAGGCTGTGAAGAGTGAAGGATGTCACCTGATTCATCATCCACGT 5632  
5633 taagcggcgccacactccagctcgtgaggtgtgagagccgaagaagacttgaatcccttaa 5692  
5633 TAAACCGCGCCACCTCCCATTTCCGGGTGGAAGGCCGAAGAACTGAAATCCCTTAA 5692  
5693 gctcgcctccatgcgtgtgttcccttacttacttaacaatactccacagggccagatgaac 5752  
5693 GCTTCGCTCCATGCGCTGCTGTTCTTCTTACTTGTCAATTAACCTCTCAGACTAATGTATGC 5752  
5753 gcttataagacgtcgaaccccatagacacttatacccttaccctgtgctatattgac 5812  
5753 GCATAGAGAGACAGCTGACCTCCATTAACCTTATCTCACCTGGTTATTAAGTACT 5812  
5813 ctgatacgggtgtcactgtataatagaactcagaggtgtgctcctagaagcactgtgtag 5872  
5813 CCGGACAGGATTTAATATATCAACACACTCAAGGGAGGCTCCTTATAGAACTCGTGCG 5872  
5873 ctgaactgcaattctgctcctcgaattgataacccgctgtttaaagcacacccccaac 5932  
5873 CTGATATATACGTTTGGCTCAGATCAGTTATTTCTTATGTCGA---CTCACCCCGAGATA 5929  
5933 taagctgtagttaggttctatctgtgtcccaagcag---agaagaagaatctgtg 5989  
5930 TCTTCATGCTCAGGATTTTATGTTTATGTTGCCACAGACCAACAATAATGMAAACATTTGG 5989  
5990 ggggtctcgtggaactcctctgtagagagatgagcgtgctacactccacagatgagact 6049  
5990 GAATCCAGAGATTTCTTTTGAACAATGGAAGCTGTATACCTTAATGATGATTT 6049  
6050 ggaatgagcagactcctccagacccggtlaaacttcccttctgca----- 6098  
6050 GGAATGGGCCAACCTCTCAGACAGATAGGTTAAGTTTCTTATGCAACACTATACA 6109  
6099 -tccggcccgaggtacaaatgaatgaactatataagaatgaagcgtcctccat 6157  
6110 GCTCTGACAAATTTAATTAATCTGACCTGATTAAGAACTGGAAGCCCAAGTCTCTCT 6159  
6158 caagctagatatactaaagataagtttcaac--gaagaagaagaagaatattcaaa 6215  
6170 CAACTTAATTAATCTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6229  
6216 agtgaataatggtatgaagctgtggaatagtttattatataatggcggg---gagacag 6271  
6230 AATGGSTAAATGGTATGCTTGGGGAATGTTATTAATTAATTAATTAATTAATTAATTA 6289  
6272 gttccacttaacactgccttgaatagatagaagcggggaagacccctgtggaatg 6331  
6290 GCTCATTTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6346  
6332 gacccgataaagctgagctgtgagaggggggggggggggggggggggggggggggggg 6391  
6347 GACCAATATACGCTTGAAGGCTCAAGACCCCAACCC----- 6385  
6392 cgggtcccaattaaactcgtcgtgagcctggaataacacagcgctagcaagtaacca 6451  
6386 -----AAGGACCAAGAACCACTCTTAACATTAATTT 6415  
6452 ctgagttgatttccctcaaacacgctagaacactcccaagttgtctgtttaagacagac 6511  
6416 CTGGATCAGACCCCACTGAGTCTACACACAGAC-----TAAATGGGGG 6460  
6512 agagacttcaagctcagagagaggttccagagccatcaactcaacccagactgtg 6571  
6461 CAAAACTTTTATGCTTATCAAGGAGGCTTTTCAAGCTTATTAATCCAGACTCCAGAGG 6520  
6572 ccacttctctgtgtgttctatcctcagggccctccattatagaaggagtgtgcta 6631



```

/db_xref="taxon:9823"
/cell_line="Shimozuma-1; malignant lymphoma derived cell
line"
/note="porcine endogenous retrovirus Tsukuba-1;
spontaneously produced virions of endogenous retrovirus
type C of swine were used to infect a swine kidney cell
line"
1..592
LTR
593..2167
/codon_start=1
/product="gag protein"
/protein_id="AAC16761.1"
/db_xref="GI:3116438"
/translation="MGQVTTPLSLTDHWTVEKSRANHLISVOKKPMQFCVSEMP
TFDVGMPSEGFENSEILAVKALIFCPGSHDOEPYILTMODLDDPPYKPMIN
KPRKGPRIALGEGKSHAEKSPHYPEIIEPPAMPPEQSVAPPYLAAGARP
PPAPGAPALVAGSPAGSRSGRACPERUTDITLPLTPTPPGQLOPLQWPPSS
ADLYNMKTNPPSEDPQRLTGLVESLMSHQTDDCQDLDTLTTEREKILIEA
RKNVPGADGRPTLQNEIDMGFLTRGWMDYNTAEGLSLIKYKALVAGLSARBP
TNLAKEVWQGNPEPSEFLERLEAFRYTTPDPDSEAKASVALAFIGQALDIR
KILAEGLQAEALRDLVKEAEKVVYKRETEERDEORKEEREERERERKREKMLT
KILAAYEGKSNTERERDFKIRSGPQSGNLGNRTPLDKQCAVCYCKEGRHARNPK
2315..5751
CDS
/note="pol protein"
/codon_start=1
/pseudo
5627..7543
/codon_start=1
/product="env protein"
/protein_id="AAC16762.1"
/db_xref="GI:3116439"
/translation="MHPITLSRHLPIRGCKPKRLKIPISFASIAWPLTSLISQTMGM
RIDDSLSNHRPLSLTLDSTGININNOGAPATGMPDLYLRLVLSISLSP
DIHAHGFYCPGPNNKHCNGNRPFCOMNCVSNQGWKWPNSODRVSFSYVN
TYNSGQFNLTWIRGSPKSPSDLDYLKISTEKGKQENILKYNWNGSWGVYTG
SKQPGSILIRIKINOLEPPAIGPNTVLTGPRTPQGGPSSNITSDDPTESSST
TKMGAKFLSLIGAFQALNSTPEATSSWCLALAPPYEGMARKNVYKEHNDQ
CWGSONKLTLEVSQKNTICKVPSHOLCHNTEAFNQSXYLVGYRMANCN
TGLTPEVSTLVNQTDFCIWQIVPRVYYPKALILDEVDYHNHOKKEPISLTAV
MLGIVAGVAGTGTAAALVTGPOOLETGLNLRITVEDIOLAEKSNLEESITSE
VVLNRRGDLILFKEGSLCVALKKECCYVYHSGAIRSMKLRRLKRRREKETT
OGMFEGRNRSPLATLILSLALTPILVLLTLVGPCTINKLAFIRERISAVQIWL
RQGYQSSPSEAR"
7548..8132
LTR
BASE COUNT 2247 a 1980 c 2030 g 1875 t
ORIGIN
Query Match 75.68; Score 6207.8; DB 4; Length 8132;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 7149; Conservative 0; Mismatches 937; Indels 139; Gaps 24;

```

```

297 tccctcagacgtaagagctgagagctgtatcaagcaagctgtctagaagatcaag 356
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TCCTCTAGGCGCGTAAGAGCTGGGGGACTGTATCAGCAGACCTGCTAGAGGATACAG 361
357 gctgcaacctggggagcgccccggaggtlvgggaagcgaaggaagctgtgtctcc 416
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 GCTGCTCCTCGGGGAGCGCCCGGAGGAGGAGAGCAGGAGGAGCGCTGGGTCTCC 421
417 tactgtcgtcagaagagcagagcttctgtgttgaagcgaagcttccccctcgagcg 476
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 TACTGTGGTCAAGAGACGAAATCTGTGTGAACCGAAAGCTTCCCTCGGACCGC 481
477 tccgactcttctcgactgtgtgaagacgcgagaggtgcggtgtgtctgagtcgtg 536
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 TCGGACTCTTTCGCTGCTGTGTGAATACGTGACGGGTCACGTGTCTGTGATCTGTTG 541
537 gttctgtcgtgtgtcttctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 GTTTCGTGTTTGTGTCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
597 acagtaactaacccccctttagttagctctcgaacatgagctgaagtaagccaggtc 656
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 ACGGTGACGACCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
657 cataatctcagtlcaagtltaagaagagaccttggcaagacttctgtcctcctaag 716
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 CATATTTTGTCACTTCAAGTTAAGAGGACCTTGCAACACTTGTGTCTCTCATAG 721
717 ccaacattgagatgtgtatgagcatalaagagggagaccttaattctgaattctctg 776
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 CCGACATTTGATGTGATGAGGCGCATCAGAGGAGCTTTAATTCAGATATCTGTGCT 781
777 gttaaagcaatacttttcaagcttgagcccgctctcctcagatcagaagagccatct 836
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 782 GTTAAAGCAATTTATTTTACAGTACGACCCGCTCTCATCTCCGATCAGAGCCCTATATC 841
837 ctacgtggaagatttggcaagaagatccctccgcatggttgaacacatggtcaataaa 896
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 842 CTACGTGCAAGATTGTGACAGGATCTCCGCCATGGGTTAAMCATGCTGATTAAG 901
897 ccaagaagaagaggttcccggaatcctgtgctcttggagagaaacaaacactcggcgaa 956
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 CCAAGAAACCCAGGTCCCGCAATCTGTGCTGTGAGAGAAAACAAACACCTGCTGAA 961
957 aaagtcgaagccctcttccctcgtatctacccgagaatgaagagcgccgagcttggcgaa 1016
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 962 AAGTCAACCCCTC-TCCTATATCTACCCGAGATGAGAACCCAGGCTTGCGCGA 1020
1017 accccaacctgttcccccaaccccttaccagcaagaggtgtgtgagagga-cctctgc 1075
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 ACCCCATGTGTTCCCGCACCCCTTATCTGTGACAGAGGATGCGGAGGAGACCTTTGC 1080
1076 cctctctgagctcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1135
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 CCTCTGTGAGGCTCGGCGGTGAGGAGCACTTCTCAAGGATCTCGACCCGAGGCGCGC 1140
1136 caaccggagggagagagagagatcgagatattacggttgcgaccttggccctccat 1195
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 CACCCGAGGCGGAGAGAGAGATGCGACATTTACCGCTGCGACAGTACGCGCTCCAC 1200
1196 gccagagggccaattgcagccctccaglatatgaccttctctcttcagatctctataa 1255
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 ACCGGGGGGCCAAATTCAGGCCCTTCAGATTTGGCCCTTTCTTCTGAGATCTCTAATA 1260
1256 ttggaataaacaatcccccttctcggagagatcccccaagcctcaggggttggtaga 1315
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 TTGGAATAACTAACCTCCCTTCTCTCGAGAGATCCCAAGCGCTCACGGGGGTGTGGA 1320
1316 gtccttctgtctctcagagcagcttggatgagatgtataaagctgtcgcagaaact 1375
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 GTCCCTTATGTTCTCTACACGCTTACTTGGATGATTTGTAAACAGCTGCTGACAGACT 1380
1376 ctccacaacggagagagagagagatctgttagagctagaanaaatgttctctgggc 1435

```



|||||  
Db 1391 CTTACACACCGAGAGGAGAGAGATTTCTATTAGAGGCTTAGAAAAATGTTCTTGGGCG 1440  
Qy 1436 cgaagcgagaccacagcagatgacaaatgagatgacatggatcttcccttgatcgcc 1495  
Db 1441 CGACGGGCGACCCACGCGGTTGCAAAATGAGATTGACATGGATTTCCTTAACCTGCGCC 1500  
Qy 1496 cggttggactacaacacgctgagatgagagagcttgaatctatcgccagctct 1555  
Db 1501 CGGTTGGGACTACAAACAGCGCTGAGAGGAGAGGCTTGAATAATCATGCGCCAGGCTCT 1560  
Qy 1556 ggtggcggtctccggggcgccctcaaaagccacctaattggtctaaagttaagagct 1615  
Db 1561 GGTGGCGGGCTCTCGGGGCGCTCAAGCGGCCACTAATTTGGCTTAAGGTAAGAGAGT 1620  
Qy 1616 gatgaagagaccgaacgaacctccctcgtaattcttgagagctcatalgaaagcctcag 1675  
Db 1621 GATGAGGAGCGAATGAAACCCCTCTGTTTCTTGAGAGGCTCTTGAAAGGCTTACAG 1680  
Qy 1676 gcggttcaaccccttgaatctactcaagagcccaagaagcctcagttggccctggcct 1735  
Db 1681 GCGGTACACCCCTTTGATCCACCTCAGAGGCCCAAAACCTCAGTGGCTTTGGCTT 1740  
Qy 1736 catggcgagtcgagctcgtgatalcagaagaacactcagagactggaaggttacaaga 1795  
Db 1741 TATAGACAGTCAGCCTTGATATTAGAAAGAGCTTCAAGACTGGAAGGGTTACAGGA 1800  
Qy 1796 ggtcagttacgtgatactagtgagagagcgagaaaggtgatacagaagagagacaga 1855  
Db 1801 GGCTGAGTTACGTGATCTAGTGAAGAGAGAGAGAAAGTAAATTAACAAAGGAGACGA 1860  
Qy 1856 aagagaagaagacagagaaagaagaagaagaagaagaagaagaagaagaagctgatag 1915  
Db 1861 AAGAAAGAGGAGACAAAGAAAGAGAGAGAAAGAGAGAGAAAGGAGAAACGTAAATA 1920  
Qy 1916 acggaagaagaagaatctgactaagatctgycgcgagtggttgaaggaagagcag 1975  
Db 1921 ACGGCAAGAGAGATTTGACTTAAGATCTTGCTGAGTGGTGAAGGAAAGCAATAC 1980  
Qy 1976 ggaagaagaagaagatcttgaagaatctagatcagccctgacagctgaaggaagctgg 2035  
Db 1981 GGAAGAGAGAGAGATTTTGAAGAAATTAAGTCAAGGCCCTTAACAGTGAAGGAACTGGG 2040  
Qy 2036 caatagagaccactcagacagagcagtgctgcatgtlaagaagaagacacatgggc 2095  
Db 2041 CATAGAGACCCACTCTGACAAAGCAATGTCATATTGTAAGAAAGAGACACTGGGC 2100  
Qy 2096 aaggaactgccccagaagaagaacaaagacgaagtcctagcttagaagaagataa 2155  
Db 2101 AAGGAACCTGCCCAAGAAAGGAAACAAAGACCAAGGATCCTAGCTTAGAAGAGATAA 2160  
Qy 2156 agattagggagagacggggtctggagccctcccgagcccaaggttaacttgaaggtga 2215  
Db 2161 AGATTAGGGGAGAGCGGGTTCCGACCCCTCCCGAGCCCAAGGTAACTTTTAAGGTGA 2220  
Qy 2216 ggggcaacagatgagatctcgtgtgatacaggaagcgagcaatcagtgctacaacc 2275  
Db 2221 GGGGCAACAGTGTGATCTCTGGTTGATACCGAGAGAAACATTCAGTCTACTACAGCC 2280  
Qy 2276 attaggaanaactaaagaanaaaatccctgggtgtagtggccacagggcaacgagta 2335  
Db 2281 ATTAGGAAAACCTAAAGATTAATAAAATCTGGGTGATGGGTG-CACGAGGCAACACAGTA 2339  
Qy 2336 tcatagactaccgaagaacgctgacttgagtgagtgagacgggtaaccacatcttct 2395  
Db 2340 TCCATGACTACCCGAAAG-ACAGTTGACTTTGGAGTGGAGCGGTAACCCACTCGTTCT 2398  
Qy 2396 ggtcattccctgagtcgccagttaccctctcagtagaagaactaactgaacaaatggagc 2455  
Db 2399 GGTCAATACCTGAGTGCACGACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGAGCA 2458  
Qy 2456 tcaatttcttggaaagaagagacagaaggtgctgtgaaataaaccacacacacgt 2515  
|||||

Db 2459 ACNAATTTCTTTGAACAAGGAGAAACGAGAGTGTCTGCATAATTAACAACTATCACTGT 2518  
Qy 2516 gttgaccctcccaatagatgatatatagactatattccccaagtaaacctgata 2575  
Db 2519 GTTGACCTCTCAATTAGATGACAAATATGACTATATCTCTCCCTAGTAACCTGATCA 2578  
Qy 2576 agatataagctcgtgtgagagacttccccaagccttgggacagaacccgaggaatgg 2635  
Db 2579 AANTATACAAATTCTGGTTGGAACAGTTTCCCAAGCCTGGGCGAAGACCGAGGATGGG 2638  
Qy 2636 ttggcaagaagaatcccccacagttatcaactgaagggcaggtgataccagatc 2695  
Db 2639 TTTGGCAAAAGCAATTCCCCACAGATTAATCACTGAAGGCGACAGTCAACAGTGTG 2698  
Qy 2696 agtcaagaataccctctgagtagaagagctcagaagaagaatctggcgcagttcaaaag 2755  
Db 2699 AGTACAGCAATACCCCTTGAGTAAAGAGCTCAAGAAAGGAATTCGGGCCGATGTCCAAAG 2758  
Qy 2756 attaatccaaagagcatcctagttccctgtccaaatcccttgaatactccctgtacc 2815  
Db 2759 ATTATACAAACAGGAGATCTAGTCTCTGTCATCTCCCTGGAAATACCTCCTGCTACC 2818  
Qy 2816 ggttagaagcctgggagccaaatgatatcgaacgaagagagacttgaagagttcaataa 2875  
Db 2819 GGTAGAAACCTGGGACTAATGACTATGACAGCAAGTACAGACTTGAGAGAGTCAATTA 2878  
Qy 2876 aaggtgacagacatacaccacacagctccgaaccccttaaacctcttgagcgccctcc 2935  
Db 2879 ACGGCTCAGAGATATACCCCAACAGTCCGGAACCTTAAACCTGTCGTCCTG 2938  
Qy 2936 gctcgaaggaactggtacacaglatggaacttaaaagtgccttcttcctgagat 2995  
Db 2939 ACCCAACAGAGAGCTGATACAGTATGACTTAAGAAAGAGCTTCTTCTGCTGAGAT 2998  
Qy 2996 acacccactagcccaacacacttcttgctcgaatgagagatccaggtacggaagaac 3055  
Db 2999 ACACCCCACTACCAACCACTTTTGGCTTCGATGAGAGATCAAGTACGAGGAAAGAAC 3058  
Qy 3056 cgggcaactcactgagaccgagactgcccgaaggttcaagaactcccgacacattga 3115  
Db 3059 CGGCACTACCTACCTGACCCGAGCTCCCAAGGTTTCAAAAGCTCCCGACCACTTTTGA 3118  
Qy 3116 cgaagccctacacaggaactcgtgccaacttcagatccaacacactcaagttgacccctc 3175  
Db 3119 CGAAGCCCTACACAGAACCTGGCACTTACAGATCAACACCTTAGGTGACCTCCT 3178  
Qy 3176 ccagttacgtgatactgctctcgtcgagagcgaacaaacaggaactccttagaaggtac 3235  
Db 3179 CCAGTACGTGATACCTGCTTGGGGGAGCCACCAACAGGACTGCTTAGAAGGACAC 3238  
Qy 3236 gaaggaactactgctggaattgtctgaactagctacagagcctcgtctaagaagccca 3295  
Db 3239 GAAGGACATCTGCTGGAATGTCTGACCTVAGGTACAGAGACCTCTCTAAGAGGGCCA 3298  
Qy 3296 gatttgcagaggaaggaataacttggggttaagtttggggcgagcagatgtgc 3355  
Db 3299 GATTTCGAGGAGAGGATACATCTTGGGTACAGTTTACGGGACGGGACGATGGCT 3358  
Qy 3356 gaaggaagcaggaagaanaactgtatgctacagataccgggccccaaacgaagcaagt 3415  
Db 3359 GACGAGGCGACGGAAGAAACGTATGTCAGATACCGGCCCAACACAGCAACAAAT 3418  
Qy 3416 gagaagtttcttgggagacgtggaatttgaagactggtggttcccggttcttgacct 3475  
Db 3419 GAGAGAGTTTGGGAGACGCTGATTTTGAAGACTGTGGAATCCGGGGTTTGGGACTT 3478  
Qy 3476 agcagcccaactctaccccgtaacaaagaagaagggagttctctcgtggctcgtgaca 3535  
Db 3479 AGCAGCCCACTACCTACCGGCTAACCAAGAAAGGGAATTTCTGAGCTCTGAGACA 3538  
Qy 3536 ccagaagcaattgtagtctatcaaaagagccctgctgagcgagcctgctcgtggccctcc 3595  
Db 3539 CCAGAAGCAATTTGATGCTATCAAAAGGCCCTGCTGAGCGCACCTGCTGCGCCCTCC 3598  
|||||



Db 5758 GCGATAGAGACAGCCGTAACCTCCATTAACCTTATCTCTACCTGGTTAATTACTGA 5817  
Qy 5811 cccctataggggtctcaactgtaaaatagaactcagagtgctgcttagagagcactgtg 5870  
Db 5818 CTCGGGACAGAGTATTAATATCAACACACTCAAGGGAGGCTCTCTTTAGAACCTGTG 5877  
Qy 5871 gcttgaactgtaattctgctcccgatctgataaccgctgtgttaaagacacactccca 5930  
Db 5878 GCGTATCTATACGTTGGCTCAGATCATGATTTCTAGTCTGAC---CTCACCCCCAGA 5934  
Qy 5931 cctagtcgtatgatatggtgtctatgtgctcagagcagag---agaagaagaactgt 5987  
Db 5935 TATCCATCTCATCGATGATTTTATGTTGGCCAGGACCCACCAATATATGAAAAACTTG 5994  
Qy 5988 tgggggtctcgggaactcctctctgaagaga tggagctgcgtcaactcccaaga tggaga 6047  
Db 5995 CGGAATCCAGAGATTTCTTTTGTAAACATGGAATGGAATGTAACCTTATGATGATA 6054  
Qy 6048 ctggaatggccgactctctccagagccggtaaaatctcccttgtaattccgccc 6107  
Db 6055 TTGGAAATGGCCACCTCTCAGCAGATAGGGTAAAGTTTCTTATGTCAACACCTATAC 6114  
Qy 6108 gggc-----aaatcaaaatgatgaactatataaagaagaagctgtccccc 6155  
Db 6115 CAGCTCTGACACATTTAATTATCTGACCTGGATTAGACTGGAAGCCCAAGTCTCTCC 6174  
Qy 6156 atcgaactagatatactaaagataagttcaact--gaaggaagaacagaataatca 6213  
Db 6175 TTCAACACCTAGATTAACATAAAATTAAGTTTCACTGAGAAAGAAAGAAATATCT 6234  
Qy 6214 aaagtggaataatgtagtagtgagtgaggagatgtttatataatagggg---ggagc 6269  
Db 6235 AAAATGGGTAATGGATGTCTTTGGGATGTATATTTATGAGGCTGGGTAACAC 6294  
Qy 6270 aggttgaacttaaacatctgccttagagatagagacgggagacacccctgtgcaat 6329  
Db 6295 AGGCTCCATTCTTAACATTTGCGCTCAAAATAAACAGCTG---GAGCTCCCAATGTGAT 6351  
Qy 6330 gggagccgaataaagtagctgtgacagagggcccccggccttgagccacgtataact 6389  
Db 6352 AGGACCAAAATACGGTCTTGACGGGTCAAAAGACCCCAACCC----- 6392  
Qy 6390 gccggggcccaactaacctcgtcgcgcgtgacataacacagccgctagcaacagtc 6449  
Db 6393 -----AAGSACCAAGACCATCTCTTAACATAAC 6420  
Qy 6450 caatgattgattcctaccacaagcctagaactcccaagtgcttctgttaagacag 6509  
Db 6421 TTCTGATCAGACCCACAGCTAGCTAGCAGCA-----CGACTAAATATGGG 6465  
Qy 6510 acagagacctcagctctatccagagagcttccaagcacaactccacagcagctga 6569  
Db 6466 GGGCAAACTTTTATGCTCATCCAGGAGCTTTTCAAGCTTTAATCTCCACAGACTCCAA 6525  
Qy 6570 tggcacttctctgtgtgctgtgtctatcccaagggcctcttataatgaggggatggc 6629  
Db 6526 GGGTACTCTCTTGTGTGGCTATGCTTATGCTTTGGGCCACCTTACTATGAAGATGGC 6585  
Qy 6630 taaagaagaanaatcaatgtgacaaagacatagaanaatcaatgtacatgggggtccg 6689  
Db 6586 TAGAAGAGGGAATTAATGATGACAAAGAACATAGAGCAATGACATGGGGATGCCA 6645  
Qy 6690 aaataagcttaccctcaactgaagtccgggaagggagacatcagaggaagaagctccc 6749  
Db 6646 AAATAAGCTTACCTTACTGAGGTTTCTGAAAAGGACACTGCATAGGAAGGTTCCCC 6705  
Qy 6750 atcccaacaacacttgcctatagtagtctgtgttataagcagggcctcagaataatcagta 6809  
Db 6706 ATCCCAACCAACCTTTGTATACCACTGAAAGCTTTAATCAAACTCTGAAAAGTCAATA 6765  
Qy 6810 tttagtaactgtgtataacaggtgtgggcatgcaataactgggttaacccctgtgttc 6869  
Db 6766 TCTGTACTCTGGTTATGACAGGNGTGGGCACTGTAATTAAGGATTAACCCCTGTGTTC 6825  
Qy 6870 caactagcttcaaccaatccaagattctgtgcatcgtgcccgaatcccccaggt 6929  
Db 6826 CACCTTGGTTTAAACCAATTAAGATTTTGTGATTAATGTCATTAATGTTCCCGACT 6885  
Qy 6930 gtaactaacctcgaagagtggtcttgaatagataatgactcgtgtataacaccaa 6989  
Db 6886 GTATTACTATCCGAAAACCAATCCTTATGATGATATGACTACGAATATCATGACAAA 6945  
Qy 6990 aagagaaccgtaacccttaaccctagctglaatgctcgtgattaggaagggcgtgtg 7049  
Db 6946 GAGAAACCCATATCTCTGACACTGCTGATGCTCGAGACTTGAGTGGAGTGGCAGAGTGT 7005  
Qy 7050 aggaacaagagacgctgcccctgaatcagagacacagacagctagagaagaagactgtga 7109  
Db 7006 AGGAACGAGACAGCTGCTGCTGTCAGGGACACAGCACTTGAANAACGACTTATGTA 7065  
Qy 7110 gtaactgctgcaatgacaagaatctccgagccttaaaagagctgtgttagcaactaga 7169  
Db 7066 CTAACATGCAATTTGTACAGAGATCTCCAAAGCCCTAGAAAATCTGTCAAGTAACTGGA 7125  
Qy 7170 agagtcctgactctctgtcgaagtgtctacagaaacggagggatagatctgct 7229  
Db 7126 GGAATCCCTAACCTCTTATCTGAAGTGTCTTACAGAAATAGAGAGGTTATATTTATT 7185  
Qy 7230 gttctaaagaaggtgtgtatgtgtcagccttaaaagaagaatgtgtctatgtaga 7289  
Db 7186 ATTTCTTAAGAAAGAGAGCATTTATGTATGCTTGAAGAGAGATGCTGTTTATGTGA 7245  
Qy 7290 tcaactgagacacacagagactcagaaagcttgaagaagcttgaagaaggtgtgt 7349  
Db 7246 TCATTTCAGGGGCCATCAGACACTCAGAACCAACTTAAAGAAAGGTTGAAGAGCTCG 7305  
Qy 7350 aagggaagagagctgacagaggggtgtgttgaagaatgtgttcaacagctcctgtgat 7409  
Db 7306 AAGGAAAAGGAAACTCTCAAGGGGTGTTGAAGGATGTTCAACAGCTCTCTTGTT 7365  
Qy 7410 gaccacctgtcttctgtcctcgaagggccctagtagtctgtcctgttactacagt 7469  
Db 7366 GGTATACCTACTTCTGCTTAAACAGGACCCCTTAATAGTCTCTCTGTTTACACAGT 7425  
Qy 7470 tgggcttgccttaataatagattgtgtcctgttgaagaagagtggtgagcca 7529  
Db 7426 TGGCCATGATTAATTAACAGTTAATGCTTATTAAGAAAGAAATAGTGAAGTCA 7485  
Qy 7530 gataatgacttaggagacagtagcaagcctcgtgacaaagagagaactgcctcta 7589  
Db 7486 GATCATGTGACTTAGACACAGTACCAAGCCCTTAGC---AGGAGAGCTGGCGCTA 7542  
Qy 7590 gcttccagcttcaagatlagaaacttaacaagaagaagtggggaatgaagaatg 7649  
Db 7543 GCTTACAGATTCTAAGATTTGAACCTATTACAGAGAGAAAGTGGGAATGAAGATG 7602  
Qy 7650 aaatgtgaactaacctccagaaacccaagaatlaataaagccttaactaactccccg 7709  
Db 7603 AAAATPACCACTTAACATTAATGAAGAGCTTAAATTTGTTCTGAATTTCCAGATTTGTTCT 7662  
Qy 7710 aatacagacctgtcgtcgcagtaaatagtagaagtgacactcctatgttca 7769  
Db 7663 TATAGTAAAGATTAGTGTGTTTGTGTTTAAATATGCGGAATGAATAGGCTGTG 7722  
Qy 7770 gggcctgtactcctgctc--aagtaagaatacaggaatagtgctgac-----taatgct 7824  
Db 7723 AGTACATGTCTAGAGCATGGAACCTTCTGAACATTTTGAATTAACAGAAAAGGAGT 7782  
Qy 7825 atctgattctgttaaaactgctgaccatagagaatgtattacacatgacagccct 7884  
Db 7783 TTCTAAGCTGTGTTTACTCTTGTAAACTGTGTCGCCCATTAAGATGTTGAATGTG 7842  
Qy 7885 atgacatatactcaactcagctgtcactcgtgccagagagccacagcagatgagact 7944  
Db 7843 ATACACATATCTTGTGTGACACATGTCTCCCAACCCGAAACATGCGCAATGTGTAC 7902



QY	1366	agagcagcgagggagagagagagagtttttaggaaatttagtctagagcccttagacagtcag	2025
Db	1440	AAAGCAATACGGAAAGAGAGAGATTTTAGAAAAATTAGTCCAGGCCCTTAGACAGTCAG	1499
QY	2026	ggacacctggccaataagaccccaatcgacaagagaccagctgtgcglatctgtaagaagaaag	2085
Db	1500	GGAACTGGGCAATAGAGCCCACTCGAACAGGACCAATGTGATATTGTAAGAAAG	1559
QY	2086	gacactgggccaagaaactctgcccagaagggagaacaaagagccgaagtccttagctctag	2145
Db	1560	GACACTGGGGAAAGAACTCCCCAAGAAAGGAAACAAAGGACTGAAGGCTTACTCTGG	1619
QY	2146	aagaagaataaagattagggagagcggggtctgagccctcccgagcccgagctlaact	2205
Db	1620	AAGAACTTAAGACTTAGGAAGACGGGGTGTGGACCCCTCCGAGCCAGGGTAACTT	1679
QY	2206	tgaagtgagaggggcaaacagttgagttcctgtgtgataccggagcggagcatcagctgc	2265
Db	1680	TGAAGGTGGGGGGCAACAGTTGAGTTCTGTGGTATACGGAGCGAAACATTCACTGTC	1739
QY	2266	tgctacacacatttagaagaactcaagaagaaanaactctcgtgtgtatgtgtgtccagagc	2325
Db	1740	TACTACAGCCATTAGGAAAACTTAAGATTAATAAATCTCGGTGTATGGGTGCCAAGGC	1799
QY	2326	aacggcagiatlcatgtgactaccggaagaccgttgacttggagtgagcggglaacc	2385
Db	1800	AACAAGATATTCATGGACTACCCCAAGAAACAGTTGACTTGGAGATGGGAGGGTAAACC	1859
QY	2386	actcgtcttcgtgtcalccctcgagctgcgcgataccctctcgtgtagaacttactgaca	2445
Db	1860	ACTCGTTTCGGTCACTACGAGTGGCCAGCACCCCTCTTAGTGAACATTATTAGCA	1919
QY	2446	agatggagctccaattctcttgcacaaggaagacagcaagtgtcctgtgaatacacac	2505
Db	1920	AGATGGAGCACAAAATTTCTTTGAAACAAGGGAACCAAGATGTCTGCAATTAACAAC	1979
QY	2506	ccatcactgtgtgacccctccaattagatgatagataatgactatattctcccgaagtaa	2565
Db	1980	CTATACACTGTGTTCACCCCTCCAAATTAGATGAGCAAAATTCGACTTATCTCCCTAGTAA	2039
QY	2566	agcctgtcgaagataatacagctctcgtgttgagagatccccaagccctggagcaagaacg	2625
Db	2040	AGCCTGATCAAAATATATCAATTTCTGTTGGTGGAACAGTTTCCCAACCTGGGCAAGAACG	2099
QY	2626	caggagatgggtttgccaagaagcaagctcccccacaggtatctaacgtgaagccagtgcta	2685
Db	2100	CAGGATGGGTTTGGCAAAACAGATTCGCCCAACAAAGTTATTCACATGAAAGGCCAGTGGCA	2159
QY	2686	caccagatatagttagcagctacccctttagtagtagaagcctcgaagaagaattctggccgc	2745
Db	2160	CACCACTGTCTAGTAGACAGTACCCCTTGAGTAAAGACTTAACAGAGGAATTCGGCCGC	2219
QY	2746	atgttcaaaagattatccaaacagggacatccctagttlctcgtlccaatcccttggaaatact	2805
Db	2220	ATGTCCAAAGATTAATCAATCCAAAGGGACTCTAATTTCTGTGCCAAATCTCCCTGATTAATCTC	2279
QY	2806	ccctgtctaacgggttaagaagcctgggagccaatgatatatgcagccgaatagaagacttgagag	2865
Db	2280	CCCTGTCTAACCGGTTTAGAAAGCCTGGGCTATATACATATGACACAGTACAGGACTTGAAGAG	2339
QY	2866	aggtcaataaagaaggtgtcagaacatacaaccaacgctccgaaccttlaaactctgtga	2925
Db	2340	AGGTCAATAAAGGGGTGCAGSATATACACCCAAACAGTCCGAAACCTTTAATACHTCTTGT	2399
QY	2926	gcgcctcccgctgaacggaactgtgtacagatattgacttaagaatgcctctctc	2985
Db	2400	GTGCTCTCCACCCCAACGGAGCTGTATACAGTATTGGACTTAAAGATGCTCTTTTCT	2458
QY	2986	gacttagaattacacccaactagccaacacacttlttgcttcgatatggaagataccagta	3045
Db	2460	GCCTAGATTACACCCCACTTAGCCAAACACTTTTGTGCTTCGAATGAGATGCCAGTA	2519

QY	3046	cggaagaacccggcgagctcaacttgagaccgcctgcccgaagggttcaagaacctcccca	3105
Db	2520	CGGGAAGAACCGGGA6CTPACTGTGACCCGACTCCCCAAAGGTTCAAGACTCCCGCA	2579
QY	3106	ccactcttgcagagccctcacacagagacccttgccaacttcaagatccaacccctcag	3165
Db	2580	CCATCTTTGAGGAAGCCCTTACACAGAGACTCTGGCCAACTTCAGATCCAAACCCCTAGG	2639
QY	3166	tgaccctctccagtaagctgagatgacccgtcttcggcgggagcccaaaccaagactgct	3225
Db	2640	TGACCCTCTCCCAATAGCTGAGTGAACCTCTTCTGGCGGAGCCACCAAGAGACTGCT	2699
QY	3226	tagaagttacgaagacactactgctgtaaattgcagactaggtctacaagacccttgcta	3285
Db	2700	TAGAAAGCAGCAAGGCACTACTGCTGGAAFTGTCAGACTTAGGCTACAGACCTCTGCTA	2759
QY	3286	agaagccccagatlttcgagagagaaggtaaacatacttgggttacagtttggggcgggc	3345
Db	2760	AGAAGGCCCAATTTGACAGAGAGAGGTAACTACTTGGGGTACGTTTGGGGAGGGC	2819
QY	3346	agcagttgctcagagagagcaggaagaanaactgtagtcagataccggcccccaacacag	3405
Db	2820	AACGATGCTTACGAGGAGGCAAGGAACAAACTGTATGTCAGATTCCGGCCCCAACCAAG	2879
QY	3406	ccaacaagttgagagaglttttggggacagcttgatalttcagactgtgatacccggt	3465
Db	2880	CCAAACAAGTGTAGAGACTTTTGTGGGACAGCTGGATTTTGGCAGACTGTGGATCCCGGGT	2939
QY	3466	ttggagactttagcgccccaccttaccggttaacaaagaanaaaggagattctccctggg	3525
Db	2940	TTGGACCTTATGACAGCCCACTTACCACCTAACCAAGAAAGGGAAATTCTCTCGGG	2999
QY	3526	ctccctgagcaccaagaagcalttgatgctatacaaaagccctgcttgagcgacctgctc	3585
Db	3000	CTCCTGAGCACACAGAGGCAATTGTGATCTATCAAAAGGCCCTGCTGAGCCACTGCTC	3059
QY	3586	tgagccctccctgagcttaactaaaccttacccttatgtagtagtgtaaggagatag	3645
Db	3060	TGAGCCCTCCCTGAGCTAACTAAACCCCTTATCCCTTATGTGATGAGCGTAAAGGAGTAG	3119
QY	3646	cccggagaggtttaaaccnaacccttagagacatlgagagagacctgttgcataactgtcoa	3705
Db	3120	CCCCGGGAGTTTAAACCAAACTCTAGACATAGGAGGAGACTGTGTGCTACCTGTCAA	3179
QY	3706	agaagcttgatcccttgtaagccagttggttgcccgatgctlcaagagctatcgagctgtg	3765
Db	3180	AGAAGCTGATCTGTAGACCGAGTGTTGGCCCTGATACCTGAAAGGCTATACGACGTGGG	3239
QY	3766	ccatctgtgttaagagacgttgacaaatltgacttltggacagagatatlaactgttaatagcc	3825
Db	3240	CCATCTGTGTCAGGACCGCTTACAAATTGACTTTGGGACAGATTAATCTGTAAATAGGCC	3299
QY	3826	cccatagtcatlgagacacatglttcggcagccccagaccogatagtatgaccaaagccgca	3885
Db	3300	CCCATGCGTTGAGAACATCTGTTGGCAGCCCCAGACCGATGAGTACCAACGCCGCCA	3359
QY	3886	tgagccactataaagcctgtcttcaagaagggttaacttltgcttcaacagccgtc	3945
Db	3360	TGAGCCCACTATCAAGACCTGTCTTCTCAAGAGAGGGTACAGCTTCCTCCACCAAGCCGCTC	3419
QY	3946	tcaacccttccacctcttcgtccttaagaagactatgaacagttgactatgattgcatc	4005
Db	3420	TCAACCGTCCCACTTCTTGCTGCTGAAGAGACTGATGAACACAGTACTCATGTATGCCATC	3479
QY	4006	aactatltgattgagagagctggggttcgcaaggaaccttaagacataacgcctgactgag	4065
Db	3480	AACTATTGATTTAGAGAGACTGGGTCGCAAGAGACCTTATAGCATAACCGCTGACTGGAG	3539
QY	4066	aagttgtaaccgtgtctcaacttgaggaagcagctatgtgtgtagaaggttaagagatgctg	4125
Db	3540	AAATGTTAACCTGTGTCACTACAGGAAAGACAGCTATGTAGTGGAAAGTAAAGAGATGGCTG	3599
QY	4126	ggcgcgagctggttgagcgggacccgcagacatcttgggccagcagcctgcccgaagaact	4185







ACCESSION	AX052634
VERSION	AX052634.1
KEYWORDS	GI:12226824
SOURCE	.
ORGANISM	porcine endogenous retrovirus. porcine endogenous retrovirus viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group. 1 (bases 1 to 7362)
REFERENCE	Federspiel, M.J.
AUTHORS	Methods to inhibit infectious agent transmission during xenotransplantation Patent: WO 00/1726-A 19 30-NOV-2000; Mayo Medical Ventures (US)
JOURNAL	Location/Qualifiers
FEATURES	1..7362
source	/organism="porcine endogenous retrovirus" /db_xref="taxon:61673"
BASE COUNT	1997 a 1621 c 1881 g 1663 t
ORIGIN	

||||| 2456 GCCTGAGTTCATGATCTAGTGAAGGAGGAGGAGAAATATATACAAAGGAGACAGA 2515  
QY 1856 agatg 1915  
Db 2516 AGAAGATTA 2575  
QY 1916 acgag 1975  
Db 2576 ACAGGCAAGATAC 2635  
QY 1976 gag 2035  
Db 2636 GGAAGATG 2695  
QY 2036 caatag 2095  
Db 2696 CAATAGATG 2755  
QY 2096 aag 2155  
Db 2756 AAGATTA 2815  
QY 2156 agatag 2215  
Db 2816 AGAGTACGATGA 2875  
QY 2216 gaggag 2275  
Db 2876 GGGGCAACAGATG 2935  
QY 2276 attag 2335  
Db 2936 ATTAGATTA 2995  
QY 2336 tcatag 2395  
Db 2996 TCATAGATG 3055  
QY 2396 ggtacatag 2455  
Db 3056 GGTACATAGATG 3115  
QY 2456 tcaaatcttcttgaag 2515  
Db 3116 ACAATTTCTTTGATTA 3175  
QY 2516 gttgacatcctaattgatgatgatgatgatgatgatgatgatgatgatgatgatgat 2575  
Db 3176 GTTACATCCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3235  
QY 2576 agatatacagtcctggttgagagagagagagagagagagagagagagagagagagag 2635  
Db 3236 AATATATCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3295  
QY 2636 ttgttgag 2695  
Db 3296 TTTGATG 3355  
QY 2696 agtgaag 2755  
Db 3356 AGTACAGATG 3415  
QY 2756 attaatcaacag 2815  
Db 3416 AATATATCAACAGATTA 3475  
QY 2816 gtttag 2875  
Db 3476 GGTAGATTA 3535  
QY 2876 aaggtgacag 2935  
|||||

Db 3536 ACAGGTCAGAGATATACACCAACAGTCCGACACCTTTATACCTTTGTGTCTCTCC 3595  
QY 2936 gcttgaagat 2995  
Db 3596 ACCCAACAGAT 3655  
QY 2996 aacacacacag 3055  
Db 3656 ACACCCACTAGAT 3715  
QY 3056 cgggag 3115  
Db 3716 GGGGATG 3775  
QY 3116 cgaag 3175  
Db 3776 GGAAGATG 3835  
QY 3176 cgaag 3235  
Db 3836 CCAGTACGATG 3895  
QY 3236 gaaag 3295  
Db 3896 GAAGGATG 3955  
QY 3296 gatctgaag 3355  
Db 3956 GATTTGATG 4015  
QY 3356 gacgag 3415  
Db 4016 GAGGATG 4075  
QY 3416 gaaag 3475  
Db 4076 GATG 4135  
QY 3476 agcag 3535  
Db 4136 AGCAGATG 4195  
QY 3536 ccaag 3595  
Db 4196 CCAGATG 4255  
QY 3596 tgaag 3655  
Db 4256 TGACGTAACAGATG 4315  
QY 3656 ttaacacacacag 3715  
Db 4316 TTTAACCCAAATCTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 4375  
QY 3716 tccctgaag 3775  
Db 4376 TCCTGATG 4435  
QY 3776 caag 3835  
Db 4436 CAAGATG 4495  
QY 3836 ggaag 3895  
Db 4496 GGAAGATG 4555  
QY 3896 tcaag 3955  
Db 4556 TCAAGATG 4615  
QY 3956 caactctgag 4015  
Db 4616 CACTCTTCTGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 4675  
|||||

QY 4016 taagagaactgggtccgaagacctaagacataccgctgactgagagaagtgtac 4075  
DB 4676 TGAGAGACTGGGTCCCGAAGACCTTACAGACATACCGCTGAGTGAAGTGTAC 4735  
QY 4076 ctgttcaactgaacgaagcagctatgtgtgaaagtaagagatgtgtggtgagc 4135  
DB 4736 CTGGTTCAGTCAAGGAGACAGCTATGTGTGAAGTAAGAGATGCTGGGGCGCGT 4735  
QY 4136 ggtgaagggagcccgcaagatctgggccaagcagctgcggaagaacttcagc 4195  
DB 4736 GGTGAGCGGAGCCCGCAGATCTGGGCGAGCCTGCGGAGGAATCTCAGCAAAA 4855  
QY 4196 ggtgaagctatgtgcccacgaagctgtggtgctggaagaatccataacat 4255  
DB 4856 GGTGAGCTATGAGGCTCCAGCAAGCTTTCGGGTGGCGGAGGGAATCTCAATACAT 4915  
QY 4256 ttatacgacagcaggtatgtccttgcgactgcacagctaacaggggcatctataa 4315  
DB 4916 TTATACAGACAGCAGATGTGCTTGGGACTGCACAGCTACAGGGGCGCATTTAAGCA 4975  
QY 4316 aaggggtgtgttactcagcaggaaggaataaagaacaagaagaattctaac 4375  
DB 4976 AAGGGGTGTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGAAATTTAAGCCT 5035  
QY 4376 attagaaccttaccattggccaagaagctagctattacacgtctctgtgacatga 4435  
DB 5036 ATTGAAGCCTTACATTTGGCAAAAGGCTATGATTAATACGCTCTTGACATCAGAA 5095  
QY 4436 agccaagaatcatalatataaggaagacagatgtgcgagcgggttgcgaagcagc 4495  
DB 5096 AGCCAAGATCCCATATCCAGAGGAACAGATGGCTGACCGGGTTCCCAAGCAGCGAC 5155  
QY 4496 ccaagctgttaactctgtcctataatagaagcccaagcccaagcccaagcagca 4555  
DB 5156 CCAGGGGTTAACCTTGCTGCTATGATGAAGAACCCCAAGCCCAAGCCCGAGACACA 5215  
QY 4556 gtacacccctagaagactgcgaagataaagaatagaccagttctctgtgactcgga 4615  
DB 5216 GTACACCTTGAAGACTGGGAGAGATGAAGAAATGAGACAGTCTCTAGACTCGGA 5275  
QY 4616 ggggaactgtctatccatataatggaagaatccctgcacacaagaaggttagaata 4675  
DB 5276 AGGAGCCTGGATACCTCAGATGGAGGAATCTCTGCCCAAAAGAGGTTAGATA 5335  
QY 4676 tgtccaacagatatacgtcttaacccaactagaactaaacacctgcagcagttgtcag 4735  
DB 5336 TGTCCAAAGATACATGCTTAACCCACCTAGGAACCTAACCTGCAGAGTTGGTACG 5395  
QY 4736 aacatccctatcatgttctgaggtacagagagtggtgactggtgtgtcaaatgtg 4795  
DB 5396 AACATCTCTTATCATGTTCTGAGGCTACCGAGAGTGGCTGATTCGTTGTAACACTG 5455  
QY 4796 tgtgcccctgcagctgttataatccttccagaatacctccaggaagaagactaag 4855  
DB 5456 TGTGCCCTGGCAGCTGGTTATGCTATCTTCCAAATACCTCCAGGAAGAGACTAAG 5515  
QY 4856 ggggaagccacccagcgtcactcgggaagtggaactcactaggtlaaagccggttaata 4915  
DB 5516 GGGGAAGCCACCCAGGCTCAGTGGGAAGGAGCTTCACTCAGAGTAAGCCGGCTAATA 5575  
QY 4916 cggaaacaatatctatgtgttctgttagacaccttccaagatgtgttagaggtctatcc 4975  
DB 5576 CGGAACAATATCTATGTGTTTTTGAACACCTTTTGAAGTGGGTGAGAGCTTATTC 5635  
QY 4976 tactaagaagaagactcaacgtgtgtgttagaagaataactcgtggaagaatttttccag 5035  
DB 5636 TACTAAGAAGAAGACTTCAACCTGGTGTAAATAAATACTGAGAGAAATTTTCCAG 5695  
QY 5036 atttgaatacctaaggtataggtgcagacaatgtgtcagcttctgttgcaggttag 5095  
DB 5696 ATTTGAATACCTAAGGTATAGGTGCAGACAATGCTTCCAGCTTTCGTTGCCAGGTAA 5755  
QY 5096 tcagggactgcccgaagatatgtgggattgtgtggaactgtgtgcatatgtcatagaccca 5155  
DB 5756 TCAGGAGCTGCCCAAGATATTGGGATTGATGTGAAGAACTGCATTTGCATACAGACCCA 5815  
QY 5156 aagctcaggaacagtaagaagagatgaataaacattaaagaacccctacccaattgtgc 5215  
DB 5816 AAGCTCAGGACAGGTAGAGAGATGAATTAAGAACCTTAAGAACCTTCAATTAATGAC 5875  
QY 5216 caacagactgtgcatlaatatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5275  
DB 5876 CGGAGAGACTGGGCTTAATGATTGATGATGCTCCGCTTGTACTTTTAAGGTTAG 5935  
QY 5276 gaacacccctggagcagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5335  
DB 5936 GAACACCCCTGGAGAGTTGGGCTGACCCCTATGATTAATCTACAGGGAGACCCCC 5995  
QY 5336 gttggcagaatattgcttgcacatagtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5395  
DB 5996 ATTGTAAGAAATTTGCTTCCGTACATAGTGTGACAGTGTGTGTGTGTGTGTGTGT 6055  
QY 5396 taagctcaagcgctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5455  
DB 6056 TAGCTCAAGGACCTTGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6115  
QY 5456 ctc---aggaggaactgtcagatccacatcgtctccaagttggaattcagctatgt 5512  
DB 6116 CTCAGGAGAGGAGACTTCCAGATCCCATCATGTTTCCAAAGGAGATTACGTACGT 6175  
QY 5513 tagagccacccgtgcaaggaacactggaactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5572  
DB 6176 TAGACCCACCGTGAGGAACTCCAGACTGCTGGAAGGGCCCTTATCAGTACTTTT 6235  
QY 5573 gaacacacccagcgt 5632  
DB 6236 GACCAACACCAAGGCTGTGAAGGTGAAGGATTTCCACTGATTCATTCATCCACGT 6295  
QY 5633 taagccggtgcacactcccgactcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5692  
DB 6296 TAAGCCGGGCGCACCTCCGATTCGGGGTGAAGCCGAAAGACTGAAGAAATCCCTTAA 6355  
QY 5693 gtttgcctcagctgt 5752  
DB 6356 GCTTGCCTCCATCGCGGT 6415  
QY 5753 gcttataagacagctcgaagccccaatagaacttcccttaactcagctgtgtgtgtgtgt 5812  
DB 6416 GCTTGTGAGACGCGCGAAGTCCCATTAACCTTATCTCAGCGTTACTTACTGACT 6475  
QY 5813 ctgatacgggtgtcactgttaatatagactcgaagtggtgtgtgtgtgtgtgtgtgtgtgt 5872  
DB 6476 CCGGTACAGGTATTAATTTAATACACTCAAGGGAGGCTCCCTGGGACCTGTTGGC 6535  
QY 5873 ctgaactgatttctgctgtcctcgtatgttaacccgcgtta-----aaagcaacctc 5926  
DB 6536 CTGAATTAATGTCTGCTTGTGATCAAGTAATCCCTGCTCAATGACCAAGCCACACCC 6595  
QY 5927 ccaactagtcgtagtatagt 5983  
DB 6596 CCGATGTACTCGCTTACGCGTTTACGTTTGTGCCAGGACCCCAATTAATTAAGAT 6655  
QY 5984 actgtggggttctgtgggaatcctcgtgttagagatgtgtgtgtgtgtgtgtgtgtgtgt 6043  
DB 6656 ATTTGGAATTCCTCAGATTTCTTTTCAAGCAATAGAGCTGGTAACTTCAATGATG 6715  
QY 6044 gagactggaatgtgcagctcctcctcgaagcaggttaaaattcctcctgt----- 6095  
DB 6716 GGAATTTGAATGTGCGCACTCTCAGCAAGCAAGATTAAGTACTTCTTTTGAACAATC 6775  
QY 6096 -----caatccggcccggtgaag 6114  
DB 6776 CTACAGTTTAATCAATTTATATGATGCAATGGAGATGGAAGATTTGGCAACGCGG 6835  
QY 6115 tacaataatgataataataaagaagactgtcctcccatcagacttatata 6174

RESULT	14
LOCUS	AF038601
DEFINITION	Sus scrofa porcine endogenous retrovirus ERV-ER15 mRNA, complete sequence.
ACCESSION	AF038601
VERSION	AF038601.1
KEYWORDS	GI:3133304
SOURCE	
ORGANISM	Pig.
REFERENCE	Eulalyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
AUTHORS	1 (bases 1 to 7333) Alyoshin,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.
TITLE	Identification of a full-length cDNA for an endogenous retrovirus of miniature swine
JOURNAL	J. Virol. 72 (5), 4503-4507 (1998)
MEDLINE	98216827
REFERENCE	2 (bases 1 to 7333)
AUTHORS	Fishman,J.A.
TITLE	Direct Submission
JOURNAL	Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REFERENCE	3 (bases 1 to 7333)
AUTHORS	Fishman,J.A.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REMARK	Sequence update by submitter
COMMENT	On May 16, 1998 this sequence version replaced gi:3116445.
FEATURES	Location/Qualifiers
SOURCE	1..7333

[illegible]

QY 361 ccaacctgggagacgccccgggaggttgaggagaccagggacgctggtgctcact 420  
|||||  
Db 371 CCACCTGGGGAGACCCCGGAGGCTGGGAGAGCCAGGAGACGCTGGTGGTCTCTACT 430  
QY 421 gtcggtcaagagaccgagtlctgtgttgaaagcgaacccctcccgccggtccg 480  
|||||  
Db 431 GTCGGGTCAAGAGACCGACTTCTGTGTGAAGCGAAGCTCCCTCCGGCGCGTCCG 490  
QY 481 actctttgctgctgtgtgaaagacgagcggtcggtggtgctgctgtgtgtt 540  
|||||  
Db 491 ACTCTTTGGCTGCTTGTGGAAGAGCGGAGCGGCTGGTGTGTCTGTGTGTGT 550  
QY 541 ctgtctgtgtgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600  
|||||  
Db 551 CTGTTTGT 610  
QY 601 tgactaccccccttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660  
|||||  
Db 611 TGACTACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 670  
QY 661 attgtcagttcaagtttaagaaagggaccttgacagacttctgtccttgaaatgcca 720  
|||||  
Db 671 ATTGTCTAGTTCAAGTTAAGAGGACCTTGACACTTGTGTGTGTGTGTGTGTGT 730  
QY 721 catctgatgt 780  
|||||  
Db 731 CATTCGATGTTGATGAGGACATCAAGGACCTTAAATTCGAATTAATTCCTGCTT 790  
QY 781 aggcacatcatlttcaagctgacccggtctcactcctgacatcaagagccctatac 840  
|||||  
Db 791 AGGCATCATTTTTCAGACTGTGACCCGCTCTCATCTGTGATCAGGACCCCTATAT 850  
QY 841 cgtgtgcaagatlttgccagaaagaccccgcaatggtttaaaccatgtgtaataa 900  
|||||  
Db 851 CGTGGCAAGATTTGGCAGAAAGATCTCCGCTATGTTAAACATGCTTAATTAAC 910  
QY 901 gaaagcaggtcccccgaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960  
|||||  
Db 911 GAAAGCCAGGTCCCGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 970  
QY 961 tcgagccctctctgtctatctaccccgagatcgaagagccgagacttgccggaacc 1020  
|||||  
Db 971 TCGAGCCCTC-TCTCGATCTACCCGAGATCGAGAGCCGCGACTTGGCCGAAACC 1029  
QY 1021 caacctgttccccaccctctacagcaaggtgtgtgtgtgtgtgtgtgtgtgt 1079  
|||||  
Db 1030 CAACCTGTCTCCACCCCTTATTCAGCAGAGGTGTGTGTGTGTGTGTGTGTGT 1089  
QY 1080 ccttgagctcgt 1139  
|||||  
Db 1090 CTTGAGACTCTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1149  
QY 1140 ccggagcggagacgagatcgcgataatcagctgtgcacactatgtgccccatgcca 1199  
|||||  
Db 1150 CCGAGGCGGAGACGAGATCGGATATACCGCTGTGCGACCTATGSCCTCCCATGCCA 1209  
QY 1200 gggggccaatttgagcccccccgatattgtgccttttctgtgcagatctataattgg 1259  
|||||  
Db 1210 GGGGGCCAATTGACGCCCCCTCACATATTGGCCCTTTTCTGTCAATCTCTAATATGG 1269  
QY 1260 aaaactlaacatcccccttctcggagagatcccccaagcctcaggggtgtgtgtgtgt 1319  
|||||  
Db 1270 AAAACTAACATCCCTCTTCTGTGGAGATCCCAACGCTCACGGGGTGTGGAGTCC 1329  
QY 1320 ctatgttctctcaacagcacttgagatgtgtcaacagctgtgtcagacactctc 1379  
|||||  
Db 1330 CTATGTCTCTCAACGCTCTACTGTGAGATTTGTCAACAGCTGTGTGACAGACTCTTC 1389  
QY 1380 acaacggagagcggagagagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1439  
|||||  
Db 1390 ACAACCGAGGAGGAGAGAAATTTCTGTGAAGGCTAAAAAAATTTCTGTGGCGCGAC 1449  
QY 1440 gggcgagcccgagctgtgcaaatgagatgacatgtgtgtgtgtgtgtgtgtgtgt 1499

Db 1450 GGGGACCCCAAGCAGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTGCCCCGGT 1509  
QY 1500 tgggtactacaacacgctgtgaagtgaggagacttgaatactatcgccagctgtgtg 1559  
|||||  
Db 1510 TGGGACTTACAAACAGGCTGAAGGTAGGAGAGCTTGAATAATCATATGTCAGCTGTGTG 1569  
QY 1560 ggggtctccggggcgctcctaagacgcccccaatttggctgaagtgaggagtgatg 1619  
|||||  
Db 1570 GCGGCTTCCGGGGCGCTTAAGACGCGCCACTTAATTTGGCTAAGGTAAAGAGGTGATG 1629  
QY 1620 caaggacggaacgaacccctcgtglatctgtgagaggtcgaatggaacccctcagcgcg 1679  
|||||  
Db 1630 CAGGAGACCAAGCAACCTCCCTCGATATTCTTGAGAGGCTCATGGAACCTTCAGCGCG 1689  
QY 1680 ttaccccttltgactactactaagagggcccaagaagcctcagtgccctgtcalt 1739  
|||||  
Db 1690 TTCACTCCCTTTTATCTTACCTTCAAGGCCCAAGACCTCATGTGCTGTGCTTCATT 1749  
QY 1740 gggcagtcggtctgtgatatcaagaaagactcagaagacttgaaaggtttacagaggt 1799  
|||||  
Db 1750 GGGCAGTCCGCTCTGTGATATCAGAGAAACTTCAAGACTGGAAGGTTTACAGAGGCT 1809  
QY 1800 gagttacgtatctatgtgagagagcagagaggtgtatcagaagggagacagagag 1859  
|||||  
Db 1810 GAGTTACGTGATCTAGTGAAGAGCAGAGAGGTATTTACAGAAAGGAGACAGAGAG 1869  
QY 1860 gagaaggaagcagagaaagaaagagagagagagagagagagagagagagagagag 1919  
|||||  
Db 1870 GAGAAGGAACAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1929  
QY 1920 caagagaagaatttgactaagatctgtgcgagtggtgtgagggagagagcagagagag 1979  
|||||  
Db 1930 CAAGAGAAAGATTTGACTTAAGATCTTGCCGCGAGTGTGTAAGGGAAGAGCAGAGAG 1989  
QY 1980 agagagagagatlttagaanaatltagltcagggccctagaagctcagggaaacctgtgcaat 2039  
|||||  
Db 1990 AGAGAGAGAGATTTTGAAGAAATTAAGTCAAGCCCTTACAGACTCAGGAACTCGGCAAT 2049  
QY 2040 aggaagccacactcgaagaaagacgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2099  
|||||  
Db 2050 AGACCCCACTGTGACAGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2109  
QY 2100 aactgcccagaagaaagaaagaaagagcagaggtcctagctcctagaagaaagataagat 2159  
|||||  
Db 2110 AACTGCCCAAGAAAGGAAACAAAGAGCCGAAGGTCTAGCTCTGAAAGATTAATAT 2169  
QY 2160 tagggagagaggggt 2219  
|||||  
Db 2170 TAGGGAGAGCGGGGTTCGAGACCCCTCCCGAGCCAGGATTAATTTGAAGGTGAGGGG 2229  
QY 2220 caaccagttgaatctcgt 2279  
|||||  
Db 2230 CAACCAAGTTGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2289  
QY 2280 ggaanaactlaaagaaanaaatactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2339  
|||||  
Db 2290 GGAANAATTAAGAAAGAAAGAAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2349  
QY 2340 tgactacccgagagaaacggttgacttgaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2399  
|||||  
Db 2350 TGACTTACCCGAGAGAACGTTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2409  
QY 2400 atccctgt 2459  
|||||  
Db 2410 ATCCCTGT 2469  
QY 2460 atttcttttgaacaaagaaagacagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2519  
|||||  
Db 2470 ATTCTTTTGAACAAGAGAAAGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2529  
QY 2520 accctcaatagatgataatgt 2579

Db 2530 ACCCTCAATTAGATGATGAATATCGACTATATTTCTCCCAAGTAAGCCTGATCAAGAT 2589  
Oy 2580 ataccgctctgttggagcaagtttcccaagccttggcaagaacgcgaaggaatggtttg 2639  
Db 2590 ATACAGTCTGTTGGACAGATTTCCCAAGCCTGGGCAAAACCCGAGGATGGGTTTG 2649  
Oy 2640 gcaaaagaagttcccccaggttatttaactgaagggcaatgtttaacagatataagtc 2699  
Db 2650 GCAAAAGCAAGTTCCCCACAGATTATTCACCTGAAGGCCAATGCTACACAGATCAAGTIC 2709  
Oy 2700 agacagtaaccccttgaagtagaagagctcgaagaagatttggccgcatgttcaagaatla 2759  
Db 2710 AACAGTACCCCTTAGTAGAGAGAGCTCGAAGAAGATTTGGCCGCAATGTTCAAAAGATTA 2769  
Oy 2760 atccacaagggatcctagtttcttccatcccttggaaatactccctgtctacaggtt 2819  
Db 2770 ATCCAAAGGCGATCTCTGATTTCTGTCAATCCCTTGAATACCTCCCTGTAACGGTT 2829  
Oy 2820 aggaagccttggagcaagatattatcgaccagtagagagcttgaagagaggtcaataaag 2879  
Db 2830 AGGAAGCCTGGAGCAATGATTATCGACAGTAGACAGCTTGAGAGAGGTCAATTAAGG 2889  
Oy 2880 gtgcagagacataccccaacggtccgaaccccttaaccccttgaagcctccgcct 2939  
Db 2890 GTGCAAGACATACCCCAAGGTCGCCGAACCTTAACTTGTAGCGCCCTCCCGCT 2949  
Oy 2940 gaacggaacttggtaacagatttgaactaaagaatgccttcttgcctgaatlaaac 2999  
Db 2950 GAACGGAACCTGTACACATGATTGACTTAAAGATGCTTCTTGTGCTGAGATTAC 3009  
Oy 3000 cccactaagcaaaccttlttgccttgcgaatggagagatccaggttgaaggaagacg 3059  
Db 3010 CCCACTAGCCAAACCTTTTACCTTCGATGAGAGATCCAGGTACGGAAGAACCGGG 3069  
Oy 3060 cagctaaccttggaccggaacttgcgcccaaggttcaagaactcccgcaacatcttgaagaa 3119  
Db 3070 CAGCTCACTGGACCCGATGCCCAAGGGTTCAAGAACTCCCGACATCTTTTGAAGAA 3129  
Oy 3120 gccctacaaggaacttgcgcaacttcaagattccaaaccctcaagttgaccttccag 3179  
Db 3130 GCCCTACACAGGAGCTGGCAACTTCAGATCCAAACCTCTGAGTGACCTCTCCAG 3189  
Oy 3180 taagtgaatgaactgtcttcttgcgaggaacccaacaggaacttgaagaagtaagaa 3239  
Db 3190 TACGTGATACCTGCTTCTGGGGGAGGCCACCAAAAGAGACTGCTTAAGGTACGAG 3249  
Oy 3240 gcaactcgtgaatgttctgaactgaagctacagaagcctctgttaagaagccagatt 3299  
Db 3250 GCACCTACTGCTGAATTTGCTGACCTAGGCTACAGAGCTTGCTAAGAAAGGCCAGATT 3309  
Oy 3300 tgcagaagagagtaataacttgggttgaagtlttgcggggcgaggaagatggttgaag 3359  
Db 3310 TGGAGGAGAGAGGTAACATATCTGGGGTACAGTTTCGGGGCGGAGCGATGGCTGACG 3369  
Oy 3360 gaagcaggaagaactgtagtcagatatccgcccacaacccaacgaagcaagatgaa 3419  
Db 3370 GAGGACGAGAAAGAACTGTAGTCCAGATACGGGCCCAACACACAGCAAAACAGTACAG 3429  
Oy 3420 gaagtttgggaagcgtgaatttgcagactgtgataccgggttgcgacctaaaga 3479  
Db 3430 GAGTTTGGGAGAGCTGATTTTGGAGACTGTGATCCCGGGTTTGGGACCTTAGCA 3489  
Oy 3480 gcccaactaccgctgaaccaaagaagggggaatttccctgggtctccttgcacccag 3539  
Db 3490 GCCCACTTACCCGCTTAACCAAGAA 3516  
Oy 3540 aagcatttgaatcatalcaaaaagccctgtgagcgacctgtccttgcctccctgac 3599  
Db 3517 3516  
Oy 3600 gtaactaaaccttacccttattgtgtgagagtaaggaagtagccgaggaagatttla 3659  
Db 3517 3516

Oy 3660 acccaaaccttagaccatgagagagacgttgcctactgtlcaagaagcttgaatcct 3719  
Db 3517 3516  
Oy 3720 gtagcagatgttgcgccgatagtctgaagctatcgcaagctgtgcatactgttcaag 3779  
Db 3517 3516  
Oy 3780 gaagctgaacaattgaacttgggaagaataataactgttaatagcccccatgtatgag 3839  
Db 3517 3516  
Oy 3840 aacatcgttcgcaagccccagaccgaatggaatgaaccaagcccgcatlgaccatlaaa 3899  
Db 3517 3516  
Oy 3900 agcctgttctcagagagaggttcaacttgcctccaccagcgctctcaacctgtccact 3959  
Db 3517 3516  
Oy 3960 ctctgcctgaagaagacttgaatgaacagtagatcatgatgttgcatacatatgatgag 4019  
Db 3517 3516  
Oy 4020 gagactggttcgcaaggaaccttaccagacataaccgttgacttggagaagtgttaacctg 4079  
Db 3517 3516  
Oy 4080 ttcaactgaaggaagcagctatgtgtgaaagttaagagatgtgttggcgagctgtgt 4139  
Db 3517 3516  
Oy 4140 gacggaccgcagatcttggccagcagccttgcggaagaagtaactlcaagcaaaagct 4199  
Db 3517 3516  
Oy 4200 gagtcatgtgcctcacaagccttgcgctgcggaaggaatacatalaacttatt 4259  
Db 3517 3516  
Oy 4260 acggaagcagatgtatgccttgcgactgcacagctacagcgggcatctataaacaag 4319  
Db 3517 3516  
Oy 4320 ggttgcctacotcagcaggaaggaataaagaacaagaagaatctlaagctatla 4379  
Db 3522 GGGTTGCTTACTTCAGAGGGAGGAATTAAGAAACAAAGAGGAATTTCTAAGCTATT 3581  
Oy 4380 gaagccttaacttgcgaagaagcctatgataataactgttctctgtgacatcagaagcc 4439  
Db 3582 GAAGCCTTAACATTTGCCAAAAAGGCTACTATTATACCTGCTGGACATCAGAAACCC 3641  
Oy 4440 aaagatcatalcttaagaagggagacagatgtgttgcggttgcggaagcagccag 4499  
Db 3642 AAAGATCTCATATTCAGAGGAGACAGATGCTGACCGGGTTGCCAAACAGCACCCAG 3701  
Oy 4500 gctgttaaccttctgcctataatagaagaagcccaagccccaagaaacccaagcagatc 4559  
Db 3702 GCTGTTAACTTCTGCTATATAGAAAGGCCCAAGGCCAGAACCCAGACAGATAC 3761  
Oy 4560 accctagaagacttgcgaagaagataaagaatagaagcaggttcttgaagctccgaaggg 4619  
Db 3762 ACCCTAGAAAGACTGGCAAGATTAAGAAAGATAGACCAAGTTCTGAGACTCCGAGGGG 3821  
Oy 4620 acctgtataactcataltggaaggaatcctgcgcccaagaagaaggttgaatattc 4679  
Db 3822 ACCTGCTTAACTCATATGGAAGAAATCTGCCCCAACAAAGAGGGTTAAGATATGTC 3881  
Oy 4680 caacagatacatcgttcaacccaactagaactaaacactgtcagcagttgttgaagaa 4739  
Db 3882 CAACAGATACATCGTCTAACCCACCTAGCACTAAACACCTGCACAGTGGTGAAGACA 3941





Db 5971 GGGCATGTATCTGGATTAACCCCTTGTCTTCCACCTTGGTTTTCACCAACCACTAAG 6030  
 Qy attctgtgcatgtgctcaaatcgtlccccagtgtaactacatccctgaagaagtgctc 6035  
 Db 6031 ACTTTGGGTATGTGTCACCAATTTGCCCGGGGTGTACTACTATCCGAAAGCAAGCATGTC 6090  
 Qy 6956 ttgatgaalatgactatcgatataacccgaacaaagaagaccgtatcccttaacctag 7015  
 Db 6091 TTGATGAATATGACTATGATATATCGGCAAAAGAGCCCATATCCCTGACACTAG 6150  
 Qy 7016 ctgtaagtgctggaattgagcgccgctgagctgaagaagagagcgccctgata 7075  
 Db 6151 CTGTAAGCTCGGATGGAGTGGCTGCAAGCGTGGCAACGAGACGGCTCCCTATATCA 6210  
 Qy 7076 caggacacacagcagtagaagaaggaacttggtagctacacatgcygccaatgaagaatc 7135  
 Db 6211 CAGGACCGCAACAGCTGGAGAAAGAGACTTAGTAACCTACATGCAATGTAAGGAGATC 6270  
 Qy 7136 tccagaccttaagaagagctgctgtaacacctagaagaagctccctgactcttctgctgaag 7195  
 Db 6271 TCCAAGCCCTAGAAATAATCTGTAGTACCTGGAGGAATCCCTTAATCTTATGTAG 6330  
 Qy 7196 tggttctacagaaacggagggatagatctgctgcttcttaagaagagtggtgtaagt 7255  
 Db 6331 TGGTTCTACAGACAGAAAGGGGTTAGTCTGTATTCTTAAGAAGAGAGGTTATGTTG 6390  
 Qy 7256 cagccttaagaagaatgtgtcttctatgtagatcactcagagagccatcagaactcca 7315  
 Db 6391 TAGCCTTAAAGAGGAATGCTGCTTCTATGTAGATCACTCAGAGCCATCGAAGATCCA 6450  
 Qy 7316 tgaacaagacttagaanaaagttagagagctgcaagaaggaagaagagctcaccagaggt 7375  
 Db 6451 TGAGCAACCTTAGAAGAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGGCTTACCAGGGGT 6510  
 Qy 7376 ggttgaagagatggttcaacagagctccttgatgtagaccacctgcttctgctcagag 7435  
 Db 6511 GGTTCGAAGAGATGTTCAACAGGTCTCCTTGATGACCACCTGCTTCTGCTCGACGG 6570  
 Qy 7436 gggccctagtagtctctgctctgttacttaacagttgggcttgccttaataatggttg 7495  
 Db 6571 GGGCCCTAGTAGTCTCTGCTCTGTACTTACAGTTGGGCGCTTGTATTAATAGTTTG 6630  
 Qy 7496 ttgcttctgttgaagaagaagatgtagtgcagctcagatcatgtagttagttaggcaacgtac 7555  
 Db 6631 TTGCTCTTGTAGAGACAGTGAATGACATTCACATCATGTACTTAGCAACAGTAC 6690  
 Qy 7556 aagcctcttgagccaagaagaactgacactcctagccttccagttctaagatagaact 7615  
 Db 6691 AAGGCTCTTGAGCCAGAGAACTGACCTTAGCCTCCAGTTCTAAGATAGTGAAC 6750  
 Qy 7616 attacaagaagaagaatgtagaagaatggaagaatggaactgaactaaccctccagagac 7675  
 Db 6751 ATTAAACAAACAAAGATGGGAAATGAAGATGAAGAAAGCAACCTCAACCTCCAGAC 6810  
 Qy 7676 ccaaggaagttaataaaacttaaaatgccccgaattacagacactgctgctgctcagt 7735  
 Db 6811 CCGAGAAAGTTAATAAAAGCTCTAATGCCCCGAATTCAGAACCTGCTGCTGCGCAGT 6870  
 Qy 7736 aataagtagaagtagtaactcctactgttctcagagcgctgctatccctgggctaaagtag 7795  
 Db 6871 AATATAGTAGAAGGTGACACTTCTATGTCTCAGGGCTGCTATCTCTGCTTAAGTAG 6930  
 Qy 7796 ataaaggaagaatgtagttagtaactgctatctgtagtctgtaaaactcagagacat 7855  
 Db 6931 ATACAGGAATAGTGTGACTAATCCCTATCTCGGATTTCTGTAAACCTACAGCGACAT 6990  
 Qy 7856 agaagaattgatacatctgacagcctagtagtgaactatcacaactgcaatctgtcaact 7915  
 Db 6991 AGAAGAAATTATACATGACAGCGCTAGTGAATCTCACTCACTCACTGCTGCTGCT 7050  
 Qy 7916 tggccagagagccagcagatgtagagctcctcagagctatcttaaaatgtagttagcagag 7975  
 Db 7051 TGCCAGAGAGCCAGATGCGGAGCTCCGGAGCTATTTTAAATGATTTGCTCCACGG 7110

Qy 7976 agcgcgagctctcgatatatttaaatgattggtccatgagcgcgagctcgtatattt 8035  
 Db 7111 AGCGGGGCTTCGATATTTTAAATGATTTGGTCCATGAGAGCGGGGCTCGATATTTT 7170  
 Qy 8036 aaatgattggttctgagcagagcttctgttgaaccataaagctcgcgatt 8095  
 Db 7171 AAATGATTGTTTGTGACGACAGGCTTGTGTGAACCCATTAAGCTGTCCGATT 7230  
 Qy 8096 cagcactggggcgagctcccttaccctctgctggtgtagagctgtagcccgagcgct 8155  
 Db 7231 CCGCACTCGGGCGCGAGTCTCTACCCCTGCGTGTGTGACAGCTGTGGGCCCGACGCG 7290  
 Qy 8156 ctggaataaataactctctgctgctgctgctatcaaaaaaa 8198  
 Db 7291 CTGGATTAATAATCTCTTGTGCTTGTGATCAATAAAAAA 7333  
 RESULT 15  
 AR130474 AR130474 7333 bp DNA PAT 16-MAY-2001  
 LOCUS AR130474 Sequence 2 from patent US 6190861.  
 DEFINITION AR130474  
 ACCESSION AR130474  
 VERSION AR130474.1 GI:14118799  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 7333)  
 AUTHORS Fishman,J.A.  
 TITLE Molecular sequences of swine retroviruses method of using  
 JOURNAL Patent: US 6190861-A 2 20-FEB-2001;  
 FEATURES location/Qualifiers  
 source 1..7333  
 BASE COUNT 1984 a 1788 c 1885 g 1676 t  
 ORIGIN  
 Query Match 66.9%; Score 5489.8; DB 6; Length 7333;  
 Best Local Similarity 83.0%; Pred. No. 0;  
 Matches 6856; Conservative 0; Mismatches 402; Indels 1005; Gaps 12;  
 Qy 1 gtagtgaagactgtggccagcgagcgttggaataaaactcctctgctgttgcac 60  
 Db 11 GTGGTGAACGACTGTGGGCCACGCGCTTGAATAAATCTCTTGTGCTTGTGATC 70  
 Qy 61 aagacgctctgctgtagttagtgggtgtgcgcttccagagcccgagcgagggat 120  
 Db 71 AAGACCGCTTCTGTGAATGATTTGGGTGTGCTCTTCCGAGCCCGAGAGGGGAT 130  
 Qy 121 tgtcttacttgagccttcaatttctgtgctgtgagcggaatcctctgcagaccctta 180  
 Db 131 TGTCTTTACTTGCGCTTTCATTTGTGTCGTGGCGGGAATCTCGACACCCCTTA 190  
 Qy 181 caccgagaaacgacttggaagtaaggaatcccttgaagactggtgtgtgtggc 240  
 Db 191 CACCGAGAACCGACTTGGAGTAAGGAGATCCCTTGGAACTATGTGTGTGCGC 250  
 Qy 241 ggcgtctgttctgagagctgtgttctggtgtagtgcgagccttgcgttgcagctgct 300  
 Db 251 GCGCTCTGTGTGAGTGTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 310  
 Qy 301 ctcaagcgtlaagacttggaagactgtgacgaagacgtgctaggaagatcacaagctg 360  
 Db 311 CTCAGACCGTAAAGACTGAGAGACTGTGATAGAGAGAGCTCTAGAGAGATCACAGCTG 370  
 Qy 361 ccacccctggggagagcccgagaggtgtaggaagagcgaagagagcgtgtgttccact 420  
 Db 371 CCACCTGGGGAGCCCGCGGAGGTGGGAGAGCCAGGAGCGCTGTGTGTCTCTACT 430  
 Qy 421 gtcggtcagagagcagcttctgtgttgaagcgaagcttccctccgagcgctcgc 480





Db 4002 CCGTCCAGCTGCTTAATGCTTAATCTCCAGAAATACCTCCAGGAAGAGACTAAGGGA 4061  
QY 4860 agccaccagcgctcacttggaaagtgaacttcaactgaagtaaaagccgctaaatacaga 4919  
Db 4062 AGCCACCCAGGCGCTCACTGGGAAGTGAGACTTCACTGAGTAAGCCGGCTTAATAGGA 4121  
QY 4920 aacaaatcatttgggttttggtaaaccttccaggaatggtagaggttactact 4979  
Db 4122 AACAAATATCTATGTTTGTGTAGACACTTTTCAGGATGGTAGAGGCTTATCCACT 4181  
QY 4980 aagaagaagactcaaccgttgggttgaagaatacttgaagaagaattttccagaatt 5039  
Db 4182 AAAAAAGAGACTTCAACCGTGTGGCTTAAGAAATATCTGGAGAAATTTTTCAGATTT 4241  
QY 5040 ggaatcactaagtaagtaaggttcagacaatgttccagcttctgtccagtaagtcag 5099  
Db 4242 GGAATACCTTAAGTAATAGGGTTCAGACAATGGTCCAGCTTTCGTTGCCAGGTAAAGCAG 4301  
QY 5100 ggaactggccaagataatgggagttgttgaactgtcattgtgatacagaacccaaagc 5159  
Db 4302 GGACTGGCCAGATATTTGGGATTTGATTGAAACTGCATTGTGCATACAGACCCCAAGC 4361  
QY 5160 tcagagaagtagagaagataagtaagaccatlaaagaagcccttaccaaattgaccaca 5219  
Db 4362 TCAGAGACAGGTAGAGAGATGATAGAAACATTAAAGAGACCTTACCAATTTGACCA 4421  
QY 5220 gagaactgcaataatgaatgtgactctcctcctgttgccttlttaagtgtagaac 5279  
Db 4422 GAGACTGGCATTAATGATGGATGGATGCTCTCTGCTTGGCTTTTAAAGGAGAGAAC 4481  
QY 5280 acccttggacaagtttgggttgcgaaccccttgaattgtctcagggggagcccccgtg 5339  
Db 4482 ACCCTGGACAGTTTGGGCTGACCCCTTAATAAATGCTTACGGGGAGACCCCGCTTG 4541  
QY 5340 gcagaattgccttgcacatagtctgctgctgtctgtccagaccttggctctagg 5399  
Db 4542 GCAGAAATTTGCTTGCACATAGTGTGATGTGCTGCTTCCAGACCTTGTGCTCTAGG 4601  
QY 5400 ctcaagcgctcgaatgggtgagcagcgatggaagcagctcgggagccactca 5459  
Db 4602 CTCAAGCGCTCGAGTGGGTGAGGACGAGCGTGAAGCAGCTCCGGGAGGCTTACTCA 4661  
QY 5460 gggaggaacttgcgaatgtccacatgcttccaaagtggagattcgaatctatgtttagcgc 5519  
Db 4662 GGAGGACACTTGCAGATTTCCACATCGCTTCAAGTTGGAATTCAGTCTATGTTAGACGC 4721  
QY 5520 caccgtcagaagaacctcgaagctcgttgaagaaggacctatctgttacttggaccaca 5579  
Db 4722 CACCGTGCAGAAACCTCGAGACTCGGTAGAAAGGACCTTATCTGTAATTTGACCA 4781  
QY 5580 ccaacggctgtgaagtcgaagaatcccaactgtgatacgtatcccaagtttagcgc 5639  
Db 4782 CCAACGGCTGTGAAGTCGANA----- 4803  
QY 5640 ggcgcacactccgactcggggttgaagaagccgaaagaactgagatacccttaagcttgc 5699  
Db 4804 -----GAATCCCTTAAAGCTTCCG 4822  
QY 5700 ctccatgcctgttcttacttacttaacaataactccccaagccagatgaagaagccttat 5759  
Db 4823 CTCATGCGGTGTCTCTTACTCTGCAATTAATCTCTCAAGTAAATGTAAGACGCTTGT 4882  
QY 5760 agacagctcgaaccccaatagacttacttacttacttacttacttacttacttacttact 5819  
Db 4883 GGAACAGCCCACTCCCAATCCCTTATCTCTCACTGTTACTTACTGACTCCGGTAC 4942  
QY 5820 ggggtgcactgtaataagactggaagtggttgccttcttaagagcacttggctggaact 5879  
Db 4943 AGGATATTAATTAACAGCACTCAAGGAGGCTCCCTTGGGAGCTGTGGCTGAAT 5002  
QY 5880 gcaatctgcctcgaattgaatacccgctgta-----aaagcaccctcccaact 5933  
Db 5003 ATATGCTGCTGCTGATCAGTATATCCCTGTCTCAATGACAGGCGCACCCCGCATGT 5062

QY 5934 agtccgaatgaatgggttctatgtcctccag-----cacagaagaagaatactgtg 5990  
Db 5063 ACTCGGTGCTTACGGGTTTTTACGTTGCCAGAGACCCCAATTAAGAAATATTGTGG 5122  
QY 5991 ggggttctgggaactcctctctgtagaagatgagctgtcgtaccctccacagatgagactg 6050  
Db 5123 AAATCCAGAGATTTCTTTTGGCAAGCATGAGGCTCATTAATCTTAATGATGGAAATG 5182  
QY 6051 gaaatggcgaatcctctcctcagagcggttaaatcttctgttgaatcc----- 6102  
Db 5183 GAATGGCCAGTCTCTCAGACAGACAGAGTAATCTTTTGTAAATCTCAACAG 5242  
QY 6103 -----ggccgggcaagtaacaa 6121  
Db 5243 TTAATCAATTAATTAATGAGCATGAGATGAGAAAGTTTGCAACAGGGGTACMAAA 5302  
QY 6122 tgatgaactataataagaagactgtctcccaactgaacttgaattatacgaataa 6181  
Db 5303 AGATG--TACGAATATAGCAAAATAGCTCATTTCTTGAACCTAGATTAATAAATAA 5361  
QY 6182 gtttactg--aaagaaacagaaatattcaaaagtgatgaatgtatgactgg 6239  
Db 5362 GTTCACTGAATAAGGAAACAGAAATATTCAAAAGGGTAAATGGTATATCTTGGG 5421  
QY 6240 gaatagttttat-----catagtcgggagcaggggtcacttcaacttgcctta 6295  
Db 5422 GAATATGATGATGAGAGGCTCTGGGAGAAACAAAGATCTGTGATATTCGCCCA 5481  
QY 6296 ggaatagaacggggagcagaacccctgttgcgaatggagacccaataagtaactgtcgaac 6355  
Db 5482 GAATGAATACTCAGATGAGACCTCCGTTGCTATGAGCAAAATAGGTTTGGCCGAG 5541  
QY 6356 aagggtccccggccttgggccaacgataactgtgcgttgcacataactgcgc 6415  
Db 5542 AAGGACTCTCAATTCAGA-----ACAGA 5565  
QY 6416 ggcctgacataacacagccgctcgaacagtaaccacttgaatttcttccaaacgc 6475  
Db 5566 GGCATCTGCTTAACCCCTCTGTATTAATTAACACTCTGATGATCCCACTGA----- 5620  
QY 6476 ctagaacactcccaagtgctctgttgaagaagcagaagacttcaactcactcaccag 6535  
Db 5621 -----GCTTAACATCACTATTAATAACAGGGGCAAACTTTTGTAGCTCATCCAGG 5670  
QY 6536 gagcttccaacccaactcaccagccctgaatgacacttcttctgttgccttgc 6595  
Db 5671 GAGCTTTTCAAGCTTAACTCACTCAAGACTCAAGAGGCTAACCTTCTTGTGGCTTGTCT 5730  
QY 6596 tatctcagggcctcttattatgaaggatgagctaaagaagaataatcaatgtgacca 6655  
Db 5731 TAGCTTGGGGCCCACTTACTATAGAGGATGGCTAGAGGAGGAAATTAATGATGACAA 5790  
QY 6656 aagagcatagaataatgataatgaggggtcccgaaataagacttaccctcaatgaagtt 6715  
Db 5791 AGGAACATAGAACCAATGATGATGGATGGATCCCAAAATTAACCTTACTAGAGTTT 5850  
QY 6716 ccgggaaggagcatgcatagaaagactccccaactcccaacacacttgcctatagta 6775  
Db 5851 CTGGAAGAGGCACTGCTAGGAGATGGTTCCCACTCCACCAACACCTTTGTAAACACA 5910  
QY 6776 ctgtgttattagcagcctcagaanaatcagattatagtaacttgcgttataaagtggt 6835  
Db 5911 CTGAAGCCTTTAATGACACTCTGAGATCAATATCTGTGCTGCTGTTATGACAGGTGT 5970  
QY 6836 gggcatgaataactgggtttaaaccctgtgttcccaactcagcttcaaccaatcaag 6885  
Db 5971 GGGCATGTAATATGATGATACCCCTTGTGTTCACCTTGTGTGTAACCAAACTTAAG 6030  
QY 6896 attctgtatagtgccaatctgtccccgaagtgactaccatcccggaagaatgttc 6955  
Db 6031 ACTTTTGGCTTATGCTCAAAATGTGCCCCGGGTGTACTACTATCCGAAAAAGCACTGC 6090









/translation="MHPTLSRRHLPTRGGEKPKRLIPISFASIAMFLTLTTPQASSK  
RLIDSSNHRPLSLWLTIDPTGVTVNSTRGVAPRGTMDELRLCLINPAKSTP  
PNLVSIFYCCPGTEKERYCGSGSEFCRWSGSDSNQDMKMPI SLODVRKSPFN  
SGPKYKMKMLYKDKSCSPDLDTLKTSTFRKSGKYSKVYKWELONSFLYGGAG  
STLTLRLIETGTEBPVAMPDVKVLAEOGPALBPNNLPPVOLTSLRPDTPPSNS  
TTGLPTNTPRNSPGVPYKTGORLPSLLOGAFOALNSIDPATISSCMICSSGPPYE  
GMAKERKENVTKERNOCOTGSRKKLITLVEYSGGTCTGKAPSPSHOHLCSTVVEQA  
SENOYLVPGYNRMWACNTGLTPCYSTVFNSKDFCVWQIVPRVYHNPBEVLEEDYD  
YRINRPKREPVSLTLAVMLGLTAVGVTGTAALITGPOLKEKGLHNAMEEDLRA  
LKEVSNLIEESTLSLEVVLONRGLDLFLREGLCAALKKECCFYVDHSGAIRDGM  
NKLRLKERRRREREDQGEFEGWENRSPMWTTLTSLUTGPIVYVLLLLLVYPCLLNR  
FVAVEREVSVAQIMVLROOYOGLLSGEMDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t  
ORIGIN

## alignment\_scores:

Quality: 2801.00 Length: 524  
Ratio: 5.345 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-171-553b-4 x A66553 ..

Align seg 1/1 to: A66553 from: 1 to: 8209

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17  
|||||  
588 ATGGGACAGACAGTACTAATCCCTTGTAGTTCGACTTCGACCATTTGGAC 637  
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34  
|||||  
638 TGAAGTTAGATCCAGGGCTCATATTTGTCAGTTAGAGTTAAGAGGGAC 687  
34 rGTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
|||||  
688 CTGGCAGACTTCTGTGCTCTGTAATGCAACATTCGAGTGTGATGAG 737  
51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAlaIle 67  
|||||  
738 CCATAGAGAGGGACCTTAAATTCGAAATATCTCGGCTGTTAAGGCAT 787  
67 eLlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84  
|||||  
788 CATTTTCAGACTGACCCGGCTCATCTGATCAGAGGCCCTATATACC 837  
84 eutThrTrpGlnAspLeuAlaGluAspProProTyrValLysProTrp 100  
|||||  
838 TTACGTGGCAGATTGGCAGAGATCTCCGCCATGAGTTAAACCATMG 887  
101 LeuAsnLysProArgLysProGlyProArgLLeuLeuAlaLeuGlyGly 117  
|||||  
888 CTAATATAACCAAGAAAGCCAGGTCGCCGAAATCTGTGCTGTGAGAGAA 937  
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134  
|||||  
938 AAACAACACTCGGCGCAAAAGTCAGCCCTCTCTCTGATCTACCCC 987  
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
|||||  
988 GAGATTCAGAGGCGCGACTGGCCGGAACCCCAACCTGTCCCCCAGC 1037  
151 ProLeuSerSerThrGlyCysGlyGlyThrSerAlaProProGlyAla 167  
|||||  
1038 CCCTTATCCACGACAGGCTGTGAGGGAGCTTGCCTCCCTCTGAGAG 1087  
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaI 184  
|||||  
1088 TCGGAGTGTGAGAGGAGCTCTCCCGGAGATCGGAGCCGAGAGGGGCA 1137  
184 hrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200  
|||||  
1138 CCCCAGAGGCGACAGAGATGCGATATTACCCCTGCGACCATTTAGC 1187

201 ProPrometProGlyGlnLeuGlnProLeuGlnIleTyrTrpProPhe 217  
|||||  
1188 CCTCCACATGCGAGGGGCCAATTCGAGCCCTCCAGTATGGCCCTTTTC 1237  
217 rSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 234  
|||||  
1238 TTCTCAGATCTCTATATTGGAAACCTAACCATCCCTTTCTCGAGAG 1287  
234 sProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250  
|||||  
1288 ATCCCAACAGCCCTCAGGGGTTGGTGAGTCCCTTATGTTCTCTCACAG 1337  
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrpG 267  
|||||  
1338 CCTACTTGGATGATGTCACACAGCTGCTGACAGACTCTTCCACACCGA 1387  
267 uGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 284  
|||||  
1388 GGAGCAGAGAGAAATTCTGTAGAGGCTAGAAAAAATGTTCTTGAGGCC 1437  
284 sPGLYArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300  
|||||  
1438 ACGGCGACCCACGACAGTTGCAAAATGAGATTGACATGGGATTTCCCTTG 1487  
301 ThrArgProGlyTrpAspTyrAsnThrAlaGlnGlyArgGlnSerLeu 317  
|||||  
1488 ACTGCCCCGGTTGGACTACACACGCGCTGAAGATGAGGAGACTTGAA 1537  
317 sLeuTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgP 334  
|||||  
1538 AATCTATCGCAGGCTGCTGCTGGCGGCTCCGGGGGCGCTCAAGACGG 1587  
334 rOThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350  
|||||  
1588 CCACATAATTGGCTAAGTAAGAGAGGTATGACAGGACCGAAGCAACCT 1637  
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367  
|||||  
1638 CCGCTGATATTCTTGAAGAGCTCATGGAAGCTTACGCGGTTCAACCC 1687  
367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 384  
|||||  
1688 TTTTGTATCTTACCTCAGAGGCCAGAAAGCTCAGTGGCCCTGCTTCA 1737  
384 LeuGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlnGly 400  
|||||  
1738 TTGGGCAATCGGCTCTGGATATCAGGAAGAACTTCAGAGCTGGAAGG 1787  
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417  
|||||  
1788 TTACAGAGAGCTGAGTTAGTGTATCTAGTACAGAGAGCGAAGAGTGT 1837  
417 rTyrArgArgGluThrGlnGlnGluLysGlnGluAlaArgLysGluVal 434  
|||||  
1838 TTACAAAGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1887  
434 rGluGluArgGluArgArgArgAspArgArgGlnGlnLysAsnLeuThr 450  
|||||  
1888 GAGAAAG 1937  
451 LysIleLeuAlaAlaValValGluLysSerSerArgGluArgGluArg 467  
|||||  
1938 AAGATCTTGGCCGCGCAAGTGTGAAGAGAGAGAGAGAGAGAGAGAG 1987  
467 gAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 484  
|||||  
1988 AGATTTTAGAAATTTAGTACAGGCCCTAGACAGTCAAGCAACTGGGCA 2037  
484 sNArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500  
|||||  
2038 ATAGACCCCACTGACAGAGCAAGTGTGCTATTGTAAAGAGAGAGAG 2087

```

501 HistripAlaArgAsnCysProLysIysGlyAsnLysGlyProLysValle 517
|||||
2088 CACTGGGCAAGAACTGCCCCAAGAAAGGAAACAAAGCCGAGGTCTCT 2137
517 uAlaLeuGluGluAspLysAsp 524
|||||
2138 AGCTTAGAAGAAATAAAGAT 2159

seq_name: gb_pat:A66552

seq_documentation_block: 8196 bp DNA PAT 29-MAR-1999
LOCUS A66552
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552.1 GI:4538105
KEYWORDS
SOURCE
ORIGIN
porcine endogenous retrovirus.
viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
source
1. 8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

alignment_scores:
Quality: 2774.00 Length: 524
Ratio: 5.304 Gaps: 0
Percent Similarity: 99.809 Percent Identity: 99.809

alignment_block:
US-09-171-553b-4 x A66552
Align seg 1/1 to: A66552 from: 1 to: 8196

1 MetGlyGlnThrValThrrhrProLeuSerLeuThrlenuAspHisrPrh 17
|||||
576 ATGGGACAGACAGTACTACCCCTTAGTTGACTCTCCACCATTTGGAC 625
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
626 TGAAGTTAGAACCCAGGGCTCATATTGTCAGTTCAAGTTAAGAGGAC 675
34 rotPrpGlnThrPheCysAlaSerLutrrProthrPheAspValGlyTrp 50
|||||
676 CTTGGCAGACTTCTGTGCTCTGAATGAGCCAAACATTGCATGTTGGATGG 725
51 ProserGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67
|||||
726 CCAATCAGAGGGACCTTAATCTGAATATATCTGCTGTAAAGCAAT 775
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84
|||||
776 CATTTTCAGACTGGACCCGCTCTCATCTCGATCAGAGAGCCCTATATCC 825
84 eutThrPrpGlnAspLeuAlaGluAspProprrProTrrPvalLysPrTrp 100
|||||
826 TTACGTGGCAGATTGTGGCAGAGATCTCTCCCATGGGTTAAACCATGG 875
101 LeuAsnLysProArgLysProGlyProArgTlleLeuAlaLeuGlyGly 117
|||||
876 CTAATTAACCAAGAAAGCCAGAGTCCCGAATCTGAGCTTGGAGAGAA 925
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrIleuProA 134
|||||
926 AAACAAACACTCGGCCGCAAAAAAGTCAGACCCCTTCTCTGATCTACCCC 975
134 rGaAspAlaGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
976 GAGATCGAGAGACCGCCGACTTGGCCGGAACCCCAACCTGTCTCCCAACC 1025
151 ProLeuSerSerThrGlyCysCysGlyGlyIleThrSerAlaProProGlyAla 167
|||||
1026 CCCTTATTCACACAGAGGTGCTGTGAGAGGACCTCTGCCCTCTCTGGAGC 1075
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaT 184
|||||
1076 TCCGCTGTGGAGGAGACCTCTGCGGAGCTCGAGCGGAGAGAGCCGCA 1125
184 hrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
|||||
1126 CCCCAGAGCGGACAGACGAGATCCGGAATATACCGCTCCGACCTATGCG 1175
201 ProprometProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSe 217
|||||
1176 CCTCCATGCCAGGGGCGCAATTCAGACCCCTCCAGTATGGCCCTTTC 1225
217 rSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 234
|||||
1226 TTCTGCAGATCTCTAATTTGMAAACTAACCATCCCTTCTCTCGGAGG 1275
234 sProGluArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1276 ATCCCAACAGCCTCAGCGGGTTGGAGTCCCTTATTTCTCTCACAG 1325
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThrGly 267
|||||
1326 CCTACTGGAGATGATGTCACACAGCTGCTGCGACACACTTCCACAAACGA 1375
267 uGluArgGluuArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaA 284
|||||
1376 GGAGCGAGAGAAATTCGTTAGAGCTTAGAAGAAATGTTCTCTGGGCGCG 1425
284 sPglYArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
|||||
1426 ACGGGGAGCCACGCGAGTTCGCAAAATGAGATGACATGGGATTTCCCTTG 1475
301 ThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGluSerLeuLys 317
|||||
1476 ACTCGCCCGGTTGGAGACTACAAACAGCGCTGAAGTAGGAGAGAGCTTGA 1525
317 sIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgP 334
|||||
1526 AATCTATGCCAGAGCTCTGTGGCGGCTCTCCGGGGCGGCTCAAGACGCG 1575
334 rOThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350
|||||
1576 CCACATATTTGGCTTAGTAGAGAGAGTGTATCAGAGGACCGAAGCAACT 1625
351 ProSerValPheLeuGluuArgLeuMetGluAlaPheArgArgPheThrPr 367
|||||
1626 CCGTCGCTATTTCTTGAAGAGGCTCATGAGAGAGCTTCAAGCGGTTACCCC 1675
367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheI 384
|||||
1676 TTTTGATCTCACTCTCAGAGGCCCAAGAAAGCTCAGTGGCCCTGAGCTTCA 1725
384 lGluGlnSerAlaLeuAspLleArgLysLysLeuGlnArgLeuGluGly 400
|||||
1726 TTGGCAGATCGGCTCTGATATCAGAGAAACCTTCAGAGACTGGAAGGG 1775
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyr 417
|||||
1776 TTACAGAGAGCTCAGTTACGTGATCTAGTAGAGAGAGCAGAGAGGTGA 1825
417 rTyrArgArgGluThrGluGluLysGluGlnArgGlyGluLysGluA 434
|||||
1826 TTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875

```

```

434  rglglugluaragglugluarargaraspargarglnolulysasnleutlr 450
      |||||
1876  GAGAAAGAAAGGAGGAGAAAGACGGTAATGACGGCCAAAGAAAGATTGTGACT 1925
451  LyslleleualalaIalValValgluglylvsSerserArggluArggluAr 467
      |||||
1926  AAGATCTTGGCCCGCAGCTGGTTGAAGGGAGAGACAGCAGGAGGAAGACAGAG 1975
467  gAspPhearglysIleargSerGlyProarglnSerGlyAsnleuGlyA 484
      |||||
1976  AGATTTTAAAGAAATTAAGGTACAGGCCCTTAGACAGTCAGGGAACCTGGGCA 2025
484  snArghrProleuasplysaspgIncyAlaTyrCysIsgluIysGly 500
      |||||
2026  ATAGGACCCCTCGACCAAGGACCAGTGTGCTATTGTMAAAGAAAAAGA 2075
501  HisTRPALaArgAsnCysProLysIysGlyAsnLysGlyProLysValle 517
      |||||
2076  CACTGGGCGAAGAACTGGCCCAAGAGAGGGAACAAAGAGCGAAGTC. CT 2124
517  uAlaleuGlugluasplysasp 524
      |||||
2125  ACCTCTAGAGAGAAAGATTAAGAT 2146

```

seq. document	block:	3025 bp	DNA	MAM	27-APR-2000
LOCUS	AF147808				
DEFINITION	Sus scrofa domestica porcine endogenous retrovirus type C gag-pol precursor, gene, partial cds.				
ACCESSION	AF147808				
VERSION	AF147808.1	GI:7650139			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
repeat_region					
primer_bind					
CDS					

[illegible]

200 yProprometProglYgInLeuGInProLeuGInTyrTrpProPheS 217  
 1631 CCTCCCATGCGGGGGCCAAATTCAGCCCTCCAGATGATGCGCCCTTTT 1680  
 217 etSerAlaSplLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233  
 1681 CTTCTGACAGATCTATATTAATGGAAACCTAACCATCCCTCTTCTCGAG 1730  
 234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG1 250  
 1731 GATCCCAACGCTCAGCGGGTGGAGAGTCCCTTATGTCCTCACCACA 1780  
 250 nPrOthrTrpAspPcysGlnGlnLeuLeuGlnThrLeuPheThrThg 267  
 1781 GCCTACTTGGAGATTTGTCAACAGCTGCTGCAAGACACTTCTCACACCG 1830  
 267 InGluArgGluArgGlyLeuLeuGluValArgLysAsnValProGlyAla 283  
 1831 AGGAGCGAGAGAGAAATTTCTTAGAGGCTAGAAAAATGTTCTCGGGCC 1880  
 284 AspGlyArgProThrGlnLeuGlnAsnGlnLysLeuAspMetGlyPheProLe 300  
 1881 GACGGGGGACCCAGCAGCTGCAAAATGAGATGACATGGGATTTCCCTT 1930  
 300 uThrArgProGlyTyrAspTyrAsnThrArgLysArgGluSerLeuL 317  
 1931 GACTCGCCCGGTTGGACTACACACGCTGAGGTAGGAGAGAGCTTGA 1980  
 317 ysiLeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 1981 AANTCTATGCCAGGCTGCTGGCGGCTCCGGGGCGGCTCACAGCGG 2030  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 2031 CCCACTAATTTGGCTAAGTAAAGAGGTGATCAGAGGACCGACGAAACC 2080  
 350 OPProSerValPheLeuGlnArgLeuMetGluAlaPheArgArgPheThrP 367  
 2081 TCCCTCGGATTTCTTACAGAGCTCAGTGAAGCTTCAGCGGCTTCACCC 2130  
 367 ropPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 2131 CTTTGTATCTACTCTCGAGGCCAGAAAGCTCAGTGGCCCTGGCCTTC 2180  
 384 IleGlyGlnSerAlaLeuAspLysArgLysLysLeuGlnArgLeuGln 400  
 2181 ATTGGCAGTCTGCTGATATCAGAAAGAACTTCAGAGACTGGAGAG 2230  
 400 yLeuGlnGluAlaGlnLeuArgAspLeuValArgGluAlaGlnLysValT 417  
 2231 GTTACAGAGAGCTGAGTTACCTGATCTAGTGAGAGAGCAGAGAGAGCTGT 2280  
 417 yrrYrArgArgGluThrGluGluLysGluGlnArgLysGluLysGlu 433  
 2281 ATTACAGAGAGGAGACAGAAAGAGAGAGAACAGAAAAAGAGAGAG 2330  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 2331 AGAGAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 2380  
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467  
 2381 TAAAGTCTTGGCCGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 2430  
 467 rGaspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 2431 GAGATTTTAGAAAAATTAGTCCAGCCCTAGACAGTCCAGGAGACCTGGCC 2480  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500  
 2481 AATAGAGCCCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2530

500 yHisTrpAlaArgAsnCysProLysGlyAlaGlnLysGlyProLysValL 517  
 2531 AACTGGGCGAGAGAACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 2580  
 517 euAlaLeuGlnGluAspLysAsp 524  
 2581 TAGCTCTAGAGAGATMAAGAT 2603  
 seq\_name: gb\_v1:PEIN33817  
 seq\_documentation\_block:  
 LOCUS PEIN33817 8849 bp DNA VRL 07-JAN-2000  
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env  
 genes and LTR (class A, clone 42).  
 ACCESSION AJ133817  
 VERSION AJ133817.1 GI:6688947  
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol  
 protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus  
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 8849)  
 Cauderna, E., Fischer, N., Bolter, K., Krach, U., Kurth, R. and  
 Toenjes, R.R.  
 Molecular Characterization of Human-tropic and  
 Replication-competent Porcine Endogenous Retroviruses  
 Unpublished  
 2 (bases 1 to 8849)  
 Toenjes, R.R.  
 Direct Submission  
 Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,  
 Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225  
 Langen, GERMANY  
 FEATURES  
 source  
 location/Qualifiers  
 1..8849  
 /organism="porcine endogenous retrovirus"  
 /proviral  
 /strain="type C"  
 /db\_xref="taxon:61673"  
 /clone="42"  
 /note="grown in PK15 (ATCC No. CCL-33) cell line"  
 1..668  
 506  
 /note="cap site"  
 671..688  
 /note="g1y4"  
 1115..2686  
 /gene="gag"  
 1115..2686  
 /gene="gag"  
 /function="retroviral capsid"  
 /product="gag protein"  
 join(1115..2686,2690..6271,6150..8132)  
 /note="viral polypeptide"  
 /codon\_start=1  
 /protein\_id="CAB65340.1"  
 /db\_xref="GI:6688948"  
 /translation="MGQTVTPPLSLTLDHMTVEVSRANLISVOYKKGPWQPFPCASEP  
 TFDVGMPSEGFENSEIILAVKAILFTQCGSHGPDQEPYILTWDLADPPVPAQAVG  
 KPRKGPRIILALGKKNHSAKVEPSRIYDEIEPTWEPGVPVPPVPAQAVG  
 PSAPGAVYVGPAPAGATRSRRGATPERTDELAFLPLRTYGPMPGGLOLPQWPFSS  
 ADLNNKTNHPFSEEDPQRLTGLVESLSLFSNQPWDDCOOLQTLFTTEBERILLBA  
 RKNVAGARGPTQONEIDMGFPIITRGWDYNTNTEGRESLKITQALVAGIRGRSRP  
 TNLAKVRVMQGPNEPSVFLERLMEAFRRTPDPYSEAKASVALAF7GOSALDIR  
 KILQRLGLOLAELRDLVREAEKYVYREDEEKEKEREERERERRORROCKNLT  
 KILAAVVGKSSRRERDFRKIRSGPQSGMLGNRPDLKDOCAVCEKGMANCPK  
 KGNKPKYALAEQKDGRRGSDPLPEPVTLLKVGQVEFLVDGASHVLLDPLGLL  
 KEKSKWVGAKQKQYQWTRRTYDIDGAGRVTHSELVIECPVPLDGRDLTKKGAQI  
 SFEQGRPEVSNNKPIYVTLQIDDERLTVSPQYKPPQDIOISMLTEOPQMAETAGMG  
 LAKQVPPQNDIOLKASAPVSVRQIPLSREAREGIMPHYQRLIQOGLIVQSPNNPDL  
 LPVRKPNNDYRPVODLREVNKRVDIHPVNPVNLISALPERRNNYTVLDLDAF  
 CLRHPTSQPLFAFEWRDPTGRTGQLTWRLPDGFKNSPTIFDEALHRLDANFKIGH

```
POYTLQVVDLLAGATKQDCLEGTALLESLSLGRASAKKAQICRREVTLYGS
LGGKRWLLEAKRTKTVQVPAITPAKQVREFLGTAGCRIMIPGATLAPLYLTFE
KEGSPMAEHOKAFDAIKALLSAPALAPDPTPEFLYDERGVAGVLOTGLGM
RRPVAIYSKLDVYASGWPVCLIAAVALIVKADKLITGONTTVIAPHALEIVVO
PPDRMTNARMTYHOSLLTERVTEAPPALINPATLLPEETDEVTADQDLIEEIG
VRKDLDTPLTGEYLTWPTDGSYYVVGKRRAGAAYVDGTTTIVASSLPSTSAQAE
LMALTOALRLAEGSINITYDSRTAFATAHVGAIVYQKGLITSAGEIKKEBLSL
LEALHLEKRLAI THCPGHOKAKDLISGNOMADAVAAQAQAVNLPIETPKAPBR
ROYTELEMOETIKIDOFSETEGTCTSYGKEILPHKEGLETVQOIRHLNHLGHIO
QVTRSPYHVLRLPGVADSVYKHCVPOLVANSPRTPGKRLGSHGAWHEVDPE
VPAKYGNKYLYLVFVDPEFSGVEAYPTKKESTYVAKIIEETEPREGICVSDNG
PAPVAOVSQGLATLIGIDMKLHCAROSSGOVBRMNTKETLTKITTTGCTINDMA
LLPVLFRVRMTPGORGLTPTPELITGSPPLVEIASVSHADVLISQLEFRLALEW
RQAKMOLREAYSEGDLOVPHRQVDSVYVRHRAGNLETRKMGYLVLLTPTAV
KVERIPTVHAISHVPPAPDPGSKAKETENPLKRLHRVVPYVNNSSMHPETLSR
HLPIRGKPKRLKIPLSFASIAMFLTLSTPOVYKRLVDPSPNHRKPLSLTLTDSG
TGININSTOGAPLGTWMPLELYVCLRSVITGLNDQAPPLVLAHYGYVYVCGPNDDE
YCNODEPFCKOMSCITSDNGNMKMPYODRYSYVNNPTSYNONTYHGRKMDQ
OROKDVRNKOISCHSIDLYLKTSTPEKROERIKQWNGISMGIVYIGSGRKKGS
VLTIRLLETOMEPPVAIGPNKGLAEGPPIQORPSPNSDYNTTSGSVTEPNTTI
KTGAKLFSLIQGAFOALNSTPPEATSSCWLCLASGPYEGMARGGFVNTKSHROC
TWSQNKLTLEVSGKGTGIGWVPSHOHCHNTEAFNRTSESOYLVPGYDRWACNT
GLPECVSTLVFNQTKDFCVWQIYPRVYYPKAVLIDKYRYRYNRKREPISITLAVM
LGLGVAGVGGTAAITGPOLBKGLSNLHRTVTEDLQLEKSYNLESLSLSEV
VLONRRGLDILFLKEGGLVAKBECFYVDHSGAIRDSKRLERLERRRRREADQ
CMPEGRNRSPPMTTLLSALNPLVLLLTIVGPCILNRFVAFVRERVAQVIMVLR
OQYGLLSQGETDL"
mat_peptide
/gene="pol"
/feature="retroviral protease and polymerase"
/product="Pol protein"
2690..6271
/gene="pol"
6150..8129
mat_peptide
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6150..8129
gene
LRR
8182..8849
8742..8747
POLYA_SIGNAL
2355 a 2191 c 2281 g 2022 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 2761.00 Length: 525
Ratio: 5.269 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.810
alignment_block:
US-09-171-553B-4 x PEN133817 ..
Align seg 1/1 to: PEN133817 from: 1 to: 8849
1 MetGlyInThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
1115 ATGGAGACGACAGTACTACCCCTTAGTTGACTCTCGACCACTTGAC 1164
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
|||||
1165 TGAAGTAGATCCAGGCGCTCATATTGTTCAGTTCAGGTTAAGAAGGAC 1214
34 rOTPrGlnThrPheCysAlaSerLutrrProThrPheAspValGlyTrp 50
|||||
1215 CTGGCAGACTTCTGTGCTCGAATAGGCCAACATTCGATGTTGGATGG 1264
51 ProSerGlnGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67
|||||
1265 CCATCAGAGGGGACCTTAATTCGAAATATATCTGCTGTAAAGCAAT 1314
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
1315 CATTTTCAGACTGAGACCGCGCTCTCATCTGATCAGGAGGCCATATATCC 1364
84 eutThrTrgGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
|||||
1365 TTACGTGGCAGATTGGCAGAAAGATCCTCCGCACTGAGTTAAACATGG 1414
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
|||||
1415 CTAAATTAACCAAGAAACCAAGCTCCCGAATCTCTGCTTGGAGAGAA 1464
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
|||||
1465 AAACAAACACTCGGCCGAAAAAGTCGAGCCCTC.TCTCGATATACCCC 1513
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
1514 GAGATCGAGGAGCGCGACTTGGCCGGAACCCCAACCTGTTCGCCACCC 1563
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167
|||||
1564 CCCTTATCCACACAGGGTGCTGTGAGGGAGACCTCTGCCCTCTGGAG 1613
167 lArProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183
|||||
1614 CTCGGGTGTGGAGGAGACTGCTGCGCGAGCTCGAGACCGGAGAGCGCC 1663
184 ThrProGluArgThrAspGlnIleAlaIleLeuProLeuArgThrArg 200
|||||
1664 ACCCGGAGCGGACGACGAGATCGCATATTCAGCTGGCAGCTATATGG 1713
200 yProProMetProGlyLysLysLeuGlnProLeuGlnTyrTrpPheS 217
|||||
1714 CCTCCCATGCGACAGGGGCCAATTCAGCCCTCGCATATGGCCCTTTT 1763
217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
|||||
1764 CTCTCGAGATCTCTATTAATTTGAAAACTAACCATCCCTTCTCGGAG 1813
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHis 250
|||||
1814 GATCCCAACAGCCTCAAGGGGTTGTGAGATCCTTATGTCTCTACCA 1863
250 nProThrTrpAspAspCysGlnLeuLeuGlnThrLeuPheThrThrG 267
|||||
1864 GCCTACTTGGATGATGATTCACACAGCTGTCGACACACTTTCACAAAG 1913
267 lGluGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
|||||
1914 AGGAGGAGAGAGAAATTTCTTAGAGGCTTAGAAAAAATGTTCTCGGGCC 1963
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
|||||
1964 GACGGGCGAGCCCGCAGTTCGCAAAATGAGATTGACATGGGATTCCTT 2013
300 uThrArgProGlyTyrAspPtyrAsnThrAlaGluGlyArgGluSerLeu 317
|||||
2014 GACTCCGCCCGGTTGGAGCTACACACAGCGCTGAGGTAGGAGAGACTTGA 2063
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyLysSerArg 333
|||||
2064 AAATCTATGCCCAGGCTCTGTGTGGGGGCTCTCCGGGGCCCTCAAGAGCG 2113
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlu 350
|||||
2114 CCACATAATTTGGCTAAGTAAGAGAGCGATGACGAGGAGCCACAGAAC 2163
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
|||||
2164 TCCCTCGGATTTCTTGAAGAGCTCATGGAACCTTCAAGGGCGTTACACC 2213
367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
2214 CTTTGTATCTTACTCTCAGAGGCCCAAGAAAGCTCAGTGGCCCTGCGCTTC 2263
```

```

384  IIEGyInSerAlaLeuAspIleArgIlystLeuGlnArgLeuGlnI 400
|||||
2264 ATTGGGCGTGGCTGTGATATCATCAAGAAACCTTCAGAGACTGGAGG 2313
400  YLeuGlnGluAgluLeuArgAspLeuValArgGluAgluIlyValT 417
|||||
2314 GTTACAGAGAGCTGAGTTAGCTGATCTAGTCAGAGAGGCAAGAAGTGT 2363
417  YTYTArGArgGluThrGluGluIlyGluIlyGluAgluIlyGlu 433
|||||
2364 ATTACAGAGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2413
434  ArgGluGluArgGluGluArgGluArgGluArgGluGluArgGlu 450
|||||
2414 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2463
450  TlySleLeuAlaIleValIleGluIlySerSerArgGluArgGlu 467
|||||
2464 TTAGATCTTGGCGGCGAGTGTGAAGGAGAGAGAGAGAGAGAGAG 2513
467  TgAspPheArgIlyIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
|||||
2514 GAGATTTTAGGAAATAGGTGAGGCGCTAGACACTAGGAGAGAGAGAG 2563
484  AsnArgThrProLeuAspIlyAspGlnCysAlaIlyIlySlyGluIly 500
|||||
2564 AATAGAGACCCCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2613
500  YHISrPAlaArgAsnGlyProIlySlyGlyAsnIlySlyGlyProIly 517
|||||
2614 ACACGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2663
517  euAlaLeuGluGluAspIlySlySlySlySlySlySlySlySlySly 524
|||||
2664 TAGCTCTAGAGAGAGAGATTAAGAT 2686
seq_name: qb_cm:AF038601
seq_documentation_block:
LOCUS AF038601 7333 bp mRNA MM 21-MAY-1998
DEFINITION Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete
sequence.
ACCESSION AF038601
VERSION AF038601.1 GI:3133304
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
pig.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE
1 (bases 1 to 7333)
Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and
Fishman,J.A.
Identification of a full-length cDNA for an endogenous retrovirus
of miniature swine
J Virol. 72 (5), 4503-4507 (1998)
TITLE
JOURNAL
MEDLINE
98216827
2 (bases 1 to 7333)
Fishman,J.A.
Direct Submission
Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
3 (bases 1 to 7333)
Fishman,J.A.
Direct Submission
Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts
General Hospital, 55 Fruit Street, Boston, MA 02114, USA
REMARK
COMMENT
On May 16, 1998 this sequence version replaced gi:3116445.
FEATURES
Location/Qualifiers
1..7333
/organism="Sus scrofa"
/strain="miniature swine"

```

```

/ab_xref="taxon:9823"
/cell_line="PK-15"
/note="porcine endogenous retrovirus ERV-PK15; Infective
virus but represents a naturally occurring transcript"
1..578
598..2172
/codon_start=1
/product="gag protein"
/protein_id="AAC16766.1"
/db_xref="GI:3116446"
/translation="MGQIVTTPSLITLDHMTVEVSRANHSVYKKGMQTECSEMP
TFDVGMPSEGEFNSLILAVAILLEQTPGSHPODEYILTWDLADPPWPVFWMLN
KPRKGPRIILALGKNSHAKVEPSRIYPEIEPPWPBPVPPYPAQAVRG
PSAPGAVVILGPGAGTSSRRGATPERTDEAILPLRTGPPMGGOLOLOYPFSS
ADLVNWKTNHPPESDPQRLTGLVSELMFESHQPTDQOLLOTLFTTEERILLEA
KKNVGAGGRTPOLONEIDMKGPPLTRPGMDVMTAGRESLKTIRQALVAGRGASRRP
TNLAKVREVMGPNPEPSPVELERLMEARRETPPTPSSEAKASVALAFITGOSLDR
KLIQLEGLQEALELDLVREAEKVYIRRETEEEKRQREKEEREERDRQENILT
KILAIVVEGKSSRERERDRFRIRSGPPOSGNIGNRTPLDKDQACAYCKEKGMWNRCPK
KNGKPKYLALEEDKD"
2320..3552
/note="truncated protein"
/codon_start=1
/product="pol protein"
/protein_id="AAC16767.1"
/db_xref="GI:3133305"
/translation="MGATGCGORQYPMWTRRTTVDLGVRTSHSEFLVPECPVLRDL
TKMGNOISPEEGRPSVNNKPIITVLTQLDDEYRLSPQYKPPDDISMTLQPOAM
AETAGMGIAKQVPPROYTQKASATPVSRQVPLSEAREGIMPHVORLIOGIIIPVQ
SFTWNPPLLEVRKPGTNDYRPVODLEVRKRODDHPTVFNPTNLISALPEERNVTYL
DKDAFPLRLHPISQPLTEFEMRDPGGRGLWTWRLPGFNKSPITFDEALHRL
ANFRDHPQVTLQYVDLLLAGATKQDCIDGTKALLLELDLGYRASAKRAQICRE
VTLYGSLRGGORMLTEARKKVTVQIPAPTTAKQVREFLTAGFCRLIIPGFATLAAP
LYPLTKRKGGLPQGGK"
4803..6725
/note="env protein; N terminus deleted"
/codon_start=1
/pseudo
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
alignment_scores:
Quality: 2758.00 Length: 525
Ratio: 5.263 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619
alignment_block:
US-09-171-553B-4 x AF038601 ..
Align seg 1/1 to: AF038601 from: 1 to: 7333
1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisrPth 17
|||||
598 ATGGAGACAGACAGTACACCCCTTAGTTGACTCTCGACCATTTGGAC 647
17 rGluValArgSerArgAlaHisAsnLeuSerValGluValIlySlyGlyP 34
|||||
648 TGAAGTTAGATCCAGGCTCATATTGTGTCACTTACAGTTAAGAGGAGAC 697
34 rOTFPGLnThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
|||||
698 CTGGCAGACTTCTGTGCTCTCTAATAGGCCAATTCGATGTGATGG 747
51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValIlyAlaIle 67
|||||
748 CCATCAGAGGAGACCTTAATTCTGAATAATTCCTGCTGTAGAGCAAT 797
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
|||||
798 CATTTTTCAGACTGAGCCGGCTCTCATCTGATCAGAGAGCCCTATATTC 847

```

84 eUThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100  
 11747 ATGGGCGAGTCGGCTCGATATACAGAAAGAACTTCAGAGACTGGAAG 1796  
 848 TTACGTGCGAAGATTGGCAGAAAGATCCCTCCCATGGGTAAACCATGG 897  
 101 LeuAsnLysProArgLysProGlyProArgLysLeuAlaLeuGlyGly 117  
 898 CTAAATAAACAAGAAAGAGCCAGCTCCCGAATCTGGCTTGGAGAGAGA 947  
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTrpLeuPro 134  
 948 AAACAACAACTCGCCGAGAAAGATCGAGCCCTC.TCCTGTATCTAACCC 996  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyTrpProThrcysSerProThr 150  
 997 GAGATCGAGAGCCCGCGACTGGCGGGAACCCCAACCTGTTCCCCACC 1046  
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167  
 1047 CCTTATCCAGCACAGGGGTGCTGTAGGGGAGACCTCTGCTCCCTGGAG 1096  
 167 lAProValValGluGlyProAlaAlaGlyTrpArgSerArgArgGlyAla 183  
 1097 CTCGGTGGTGAGGGAGCTGCTCGCGGAGTCGAGCCGAGAGAGCCGCC 1146  
 184 ThrProGluArgThrAspGluLeuAlaLeuLeuProLeuArgTrpGly 200  
 1147 ACCCGGAGCGGACAGACGATCGCATATTACCGTCGCGACACTATGG 1196  
 200 YProPrometProGlyGlyGlnLeuGlnProLeuGlnTrpTrpProPhe 217  
 1197 CCTCCCATGCCAGGGGCCAATTCCACCCCTCCAGATTGGCCCTTTT 1246  
 217 eSerAlaAspLeuTrpAsnTrpLysTrpAsnHisProProPheSerLys 233  
 1247 CTTCGCGATCTCTATTATGGAAACTAACCATCCCTCTTCGAGAG 1296  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly 250  
 1297 GATCCCAAGCGCTCAGCGGGTGGTGGAGTCCCTTATGTTCTCTCACCA 1346  
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrp 267  
 1347 GCCTACTGGGATGATGTCAACAGCTGCTCAGACACTCTTCACCAACG 1396  
 267 lGluArgGluArgLysLeuLeuGlnAlaArgLysAsnValProGlyAla 283  
 1397 AGGACCGAGAGAGATTCTGTAGAGGCTAAAAAAATGTTCTCGGGGCC 1446  
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLe 300  
 1447 GACGGGCGACCCAGCAGCTTCCAAATATGACATTGACATGGGATTTCCCT 1496  
 300 uThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgLysSerLeu 317  
 1497 GACTCGCCCGGTGGGACTACAAACAGCGCTGAAGGTAGGAGAGCTTGA 1546  
 317 ySileTrpArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 1547 AAACCTATCGCCAGGCTCTGGTGGCGGCTCTCCGGGCGCCTCAAGACG 1596  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlyPr 350  
 1597 CCCACTAATTTGGCTAAGGTAGAGAGGTGATGAGGAGCCGAGAGCAAC 1646  
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367  
 1647 TCCTCGGTATTCTTGAGAGGCTCATGGAAGCCTTCAGCGGTTCACCC 1696  
 367 rOPIeAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 1697 CTTTGTATCTTACTCAGAGGCCAGAAAGCTCAGTGGCCCTGGCTTC 1746  
 384 lIleGlyInSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400

1747 ATGGGCGAGTCGGCTCGATATACAGAAAGAACTTCAGAGACTGGAAG 1796  
 400 YLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyLysVal 417  
 1797 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAGAGAGCAGAGAGAGTGT 1846  
 417 YTrpArgArgGluThrGluGluLysGluAlaArgLysGluLysGlu 433  
 1847 ATTACAGAGAGGACACAGAGAGAGAGAGACAGAGAGAGAGAGAGAG 1896  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuThr 450  
 1897 AG 1946  
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467  
 1947 TACATCTTGCCCGCAGTGGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 1996  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 1997 GAGATTATTAGAAATATAGGTACAGCCCTTACAGAGTACAGGAACTGGC 2046  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTrpCysLysGluLysGly 500  
 2047 AATGAGACCCCACTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2096  
 500 YHisTrpAlaArgAsnProLysGlyLysAlaGlnLysGlyProLysVal 517  
 2097 ACATGGGCAAGAGACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2146  
 517 eAlaLeuGluGluAspLysAsp 524  
 2147 TAGCTTAGAAGAAAGATTAAGAT 2169

seq\_name: gb\_pat:ARI30474

seq\_documentation\_block:

LOCUS ARI30474 733 bp DNA

DEFINITION Sequence 2 from patent US 6190861.

ACCESSION ARI30474

VERSION ARI30474.1 GI:14118799

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 733)

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 2758.00

Ratio: 5.263

Percent Similarity: 99.810

Percent Identity: 99.619

alignment\_block:

US-09-171-553B-4 x ARI30474

Align seg 1/1 to: ARI30474 from: 1 to: 733

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrp 17

598 ATGGGAGACAGAGTACATCCCTTACTTGGATCTCGACATTTGAC 647

17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGly 34



648 TGAAGTAGATCCAGGCGCTCAATTTGTGACGTTCAAGGTTAAGAGGAC 697  
 34 rOTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
 698 CTTGGCAGACTTCTGTGCTCTGATGAGCCACATTCGATGATTGGAGG 747  
 51 ProSerLysGluThrPheAsnSerGluLeuLeuAlaValLysAlaI 67  
 748 CCATAGAGGGGACCTTAAATTCTGAAATTTATCCGTGGCTTTAAGCAT 797  
 67 eIIePheGlnThrGlyProGlySerHisProAspGlnLupProTyrIle 84  
 798 CATTTTTCAGACTGAGCCGCTCATTCGATCAGAGCCCTATATCC 847  
 84 eUTThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100  
 848 TTACGTGGCAGATTTGGCAGAAAGATCCTCGCCATGGGTTAAACAGC 897  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117  
 898 CTTAAATTAACCAAGAAAGCCAGGTCCTCCGAAATCCTGGCTTTGGAGAGA 947  
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134  
 948 AATCAACACCTGGCCGAAAGATCGAGCCCTC.TCCTCGTATCTACCCC 996  
 134 rAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 997 GAGATCGAGGAGCCGCCGACTTGGCCGAGACCCCAACCTGTTCCGCCAC 1046  
 151 ProLysSerSerThrGlyCysCysGluGlyThr.SerAlaProProGlyA 167  
 1047 CCTTATTCAGACAGAGGTCGCTGAGGAGACCTCTGCCCTCTCGAGAG 1096  
 167 lApProValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
 1097 CTCGGTGGTGGAGGAGACTGCTGCCGGGACTCGAGCCGAGAGAGGCC 1146  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
 1147 ACCCGGAGGAGACAGACGAGATCCGATATTACCGCTGCGACCTATGG 1196  
 200 yProPrometProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProphe 217  
 1197 CCTCCCATGCGCAGGGGCCAATTCAGGCCCTCCAGTATTTGCCCTTT 1246  
 217 erSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerG 233  
 1247 CTTCTGACAGATCTTATATATGAAACCTAACCATCCCTTTCTCGAG 1296  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250  
 1297 GATCCCAAGCCCTCAGGGGTTGGAGTCCCTTATGTTCTCTCACCA 1346  
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTr 267  
 1347 GCTACTTGGAGATGATGTCACACGCTGTCAGACACTCTCACAAACG 1396  
 267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
 1397 AGAGAGGAGAGAAATTTGTTAGAGGCTAATAAAAAATTTCTCTGGGCC 1446  
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle 300  
 1447 GACGGCGAGCCAGCAGAGTTGCAAAATGATGATGACATGGATTTCCTT 1496  
 300 uThrArgProGlyLysPheAspThrAlaGluGlyArgGluSerLeuL 317  
 1497 GACTCGCCCGGTTGGACTACACACGCGCTGAAGGTAGGAGACTTGA 1546  
 317 ystLeuTrpArgGlnAlaLeuValAlaGlyLeuArgGlyLysSerArg 333  
 1547 AATTTATCGCAGGCTGTGGTGGGCTCCGCGGCGCTCAAGACGG 1596

334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 1597 CCCACTAATTTGGCTAAGGTAACAGAGTGATCATCAGGAGCCAGACGACC 1646  
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367  
 1647 TCCCTCGTATTCTTGAGAGGCTCATGGAAGCTTCAGGCGGTTCAACC 1696  
 367 ropAspProThrThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 1697 CTTTGTATCTTACTCTCAGAGCCAGAAAGCTTCAGTGGCTTGCCCTTC 1746  
 384 lIleGlyInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400  
 1747 ATTGGCAGTCGCGCTCGCATATCAGAAAGAACTTCAGAGACTGGAGAG 1796  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyLysVal 417  
 1797 GTTACAGAGAGCTGATGATTACGTATGATGAGAGAGCAGAGAGAGGT 1846  
 417 yTyrArgArgGluThrGluGluGlyGluGluArgGlyGlyGlyGlu 433  
 1847 ATTACAGAGAGGAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1896  
 434 ArgGluGluArgGluGluArgArgArgArgGlnGluLysAsnLeuThr 450  
 1897 AGAGAAAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946  
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467  
 1947 TAAGATCTTGGCCGCTGAGTGGTGAAGGAGAGAGAGAGAGAGAGAG 1996  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 1997 GAGATTTTGAAGAAATTAAGTCAGAGCCCTGACAGGTAGAGGAACCTG 2046  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG 500  
 2047 AATAGAGACCCACTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2096  
 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517  
 2097 ACACCTGGGCAAGGAACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 2146  
 517 euAlaLeuGluGluAspLysAsp 524  
 2147 TAGCTTAGAGAGAAATTAAGAT 2169  
 seq\_name: gb\_v1: PEN133818  
 seg\_documentation block:  
 LOCUS PEN133818 8750 bp DNA VRL 14-APR-2000  
 DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env  
 genes and LTR (class B, clone 43).  
 ACCESSION AJ133818  
 VERSION AJ133818.1 GI:6688949  
 KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol  
 protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM Porcine endogenous retrovirus  
 Viruses; Retroid viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 8750)  
 Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.  
 Establishment and characterization of molecular clones of porcine  
 endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)  
 JOURNAL MEDLINE 20219394  
 REFERENCE 2 (bases 1 to 8750)  
 AUTHORS Toenjes,R.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,

Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225  
Langen, GERMANY

# FEATURES

source Location/Qualifiers  
1..8750  
/organism="porcine endogenous retrovirus"  
/proviral  
/strain="type C"  
/db\_xref="taxon:61673"  
/clone="43"  
1..629  
LTR  
misc\_feature  
/note="cap site"  
632..649  
/note="gly4"  
1064..2635  
/gene="gag"  
1064..2635  
mat\_peptide  
/gene="gag"  
/function="retroviral capsid"  
/product="Gag protein"  
join(1064..2635,2639..6223,6099..8072)  
/note="viral polyprotein"  
/protein\_id="CAB65341.1"  
/db\_xref="GI:6688950"  
/translation="MGOTVTPPLSLIDHWT EYRSRAHNL SVYKKGPWQFCASEMP  
TFDWPSEGTENSEILLAVKAILFQTPSSHDDEYILITWDLADEPPWPKPMIN  
KPRKPGRIILAEKNSAEKVEPRIRIPEEPPEPPPPPPAGAAAG  
PSAPGPAPVEGPAAGTSGRAGATPEKTDILILPLRTYPPPGOLOPQWPPS  
ADLYNMTNHPPESEDFORLGLVESLMSHOPWMDCCOLLOTFTTEREILL  
RKINPGADGRPTOLONEIDMGPILTRGMDYNABRESKITYROLVAGLSASRP  
TNLAKEVWOGNEPSPVLEKMEFRFTPTDSEKOKSAVALAETGOSALDR  
KILORLEGOELERDLVREAEVYTRKTEEEKERKEEREREERDRROKEMLT  
KILAAVEGKSSREKEREIRISGRPSGONLNTPLDKDCAYCKEKGHMAKPK  
KGNKPGVIALLEDKDRGSDPLPEPRVALKEGQVEELVDTGAHSLVLOPLKTL  
KEKSWMGATGORYWTRRTVDLAGVTRTFLVTPCEPVLIGRLITRMGOI  
SFGQREPVNKKPIYTLTLOLDEERYLSPQVKDOOIOSLEOPPMATATAGM  
LAKOVPOVTOIKASATPVSVROYLSREAREBIEHVOBLIOGLIVEVOSTWNPPL  
LPVPRGTNDYRPVODIREYKRVODIHPTVPNPMILSLAPERWYVYLDAKAF  
CLRLHPTISPLFAFEMRDPGTGRGOLTWRLPOGKNSPTIFDEALHDLNFIH  
POVTLLOYDILLAGATKODCEGKALILELSDGYRASAKAOICREVTYLGYS  
LRGGORMLTARRKTVQIAPPTAKOVBELGTAGCFRLMIPGFATLAAPLYLKE  
KGEENAPBEHOKTFDAIKKALISAPALIPDYKPEYLVDEPKGAVRVLITGLPM  
RRPVAYLSKLDIVASGWPCLKAIAVALIKYKADKILCONITVIAHALENIVRO  
PPRRMTNARMTIYOSLLTERVTPAPALNATILPEETEPVHDOHLLIETG  
VRDLDITLPGVLTWFDGSSYVEGRMGAAYVDGTHIMASLEEGSAQAE  
LMAIOLARLAEKSNITYDSRYAFATAHGAIKOGILTSAGRETKNEEILSL  
LEALHLPRKLAIIHCPGHOKAKDLISRGONOMDRVAKOAOAVNLITETPKAPPR  
ROYLEDMOEIKIIOFSEPEGTCTSYGKTELPHKBELEYOOHRTHTHGLKHO  
OLVRTSPYHVLRLPGVADSVKHCYPCQIYANPNSVPPGKRIRGSHGPAHMEVDTE  
VPRAKGNKTLIVDTESGVAEYPTKETSTVAKKILLETIFPFGIPKVIKISNG  
PAFVAQVSOGLAKILIGIDMKLHCAYRPOSSGOVERNKKIKETLRLTETINMIA  
LPLFVLRVNTPGOGLTPEYELLYGPPLEIIVESHVADVLISPLESRKALENIA  
ROKAMKOLREAYSGEDLOVHRFOVGDVYVRHAGLLETRMWSGPTVYLITTPAY  
KVEGISTWIIASHVUKLAPPDGGWRAKTEKNEILKILHLVYSNNNSGOMHPLISW  
RHLPTGGEKRLRLTSLASTAMETLITITTOASKRILDSNPRPLSLTWLITDP  
DTGVYTNSTRGVAPRGTWPELHFCRLINPAVKSTPMLVSYGKYKIDGEEKKCY  
GGSEESFCRMSVCVNSNDGMPKIPISLODRVAFSEVNSPGYKAYKIDGSCSPD  
LDYLKISPEKROENIOKIMSGWIVFYRYGGAGSTLITRIEIGETPPAVAG  
PKVLEOGRPALPEPHNLPPOLITSLRPDIOPNSNGTGLIIPNTPNSGVAVKT  
GORFSLIGAFQAINSTDPAATSSCWLCTSYGYRYEEMAEKGFNTYKERNOCTM  
GSNNKTLTLEVGKCTCKAPRSHOHCISYVYEQASENOLVAPGINRMACNTGL  
TRCVSSSVNOSKDFCVMVOIVPRVYHPEEVVLEDEYDIRNPRKREPSILTAVMG  
LGTAVGVGTGTAALITGPOLKEKGLAEHAAETEDLRALKEESVNLSELSISVIL  
ONRKGDLTLEREGGICALKECCFYVDHSGAIDSMSKLEERRERREBDEOGM  
FKGWRSPMTTLLSALTGLPLVLLLTITVGPCLINRVAIVRRERVAVOIMVLRQ  
YOGLSQGETDI"  
2639..6223  
mat\_peptide  
/gene="pol"  
/function="retroviral protease and polymerase"  
/product="Pol protein"  
2639..6223  
/gene="pol"

mat\_peptide 6099..8069  
/gene="env"  
/function="retroviral envelope"  
/product="Env protein"  
gene 6099..8069  
/gene="env"  
LTR 8122..8750  
POLYA\_signal 8643..8648  
BASE COUNT 2288 a 2214 c 2266 g 1962 t  
ORIGIN

alignment\_scores:  
Quality: 2755.00 Length: 525  
Ratio: 5.25 Gaps: 1  
Percent Similarity: 99.810 Percent Identity: 99.619

alignment\_block:  
US-09-171-553B-4 x PEN133818 ..  
Align seg 1/1 to: PEN133818 from: 1 to: 8750

1 MetGlyGlnThrValIThrThrProLeuSerLeuThrLeuAspHisTrpTh 17  
|||||  
1064 ATGGGACAGACAGTACGACGCCCTTATGTTGACTCTCCACCATTTGGAC 1113  
17 rGluValArgSerArgLanHisAsnLeuSerValGlnValLysGlyLP 34  
|||||  
1114 TGAAGTTAGATCCAGGCGCTATAATTGTCTAGTTCAGGTTAAGAGGAGC 1163  
34 rOTrPGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
|||||  
1164 CTTCGACAGACTTCTGCTCTGAATGAGCCACATTCGATGTTGATGG 1213  
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67  
|||||  
1214 CCATCAGAGGGGACCTTTAAATTCGAATAATATCTGCTGTTAAGCAAT 1263  
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
|||||  
1264 CATTTTTCAGACGTGAGCCAGCTCTCATCTGTATCAGAGCCCTATATTC 1313  
84 eUthrTrpGlnAspLeuAlaGluAspProProThrValLysProTrp 100  
|||||  
1314 TTACGTGGCAAGACTTGGCAGAAATCTCCGCGATGGGTTAAACCATGG 1363  
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyL 117  
|||||  
1364 CTAAATTAACCAAGAAAGCCAGGCTCCCGAATCTGCTTGGAGAGAA 1413  
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134  
|||||  
1414 AACCAAAACATCGCGCGAAGAAATCCAGCCCTC. TCTCTGTATCTACCC 1462  
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
|||||  
1463 GAGATGAGAGAGCGCGGCACTTGGCGGAAACCCAAACGCTGTTCCCCACC 1512  
151 ProLeuSerSerThrGlyCysGlyGlyThr.SerAlaProProGlyA 167  
|||||  
1513 CCTTATTCACGACAGAGGTCGTGTGAGGGAGCCCTGCTCCCTCGAG 1562  
167 laProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
|||||  
1563 CTCGGGTGTGGAGGAGCTGCTCGCGGACTCGGAGCGGAGAGGAGCGCC 1612  
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
|||||  
1613 ACCCGGAGCGGACAGACAGATCGGATATTAACGCTCGGACCATATGG 1662  
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217  
|||||  
1663 CCTCCCATGCGGGGGGCAATTCAGCCCTCCAGATATGCGCCCTTTT 1712

217 eSerAlaAspLeuTyrAsnTrrpAsnHisProProPheSerGlu 233  
 1713 CTTCTGAGATCTATATGAAACTAACCATCCCTTTCTCGAG 1762  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG1 250  
 1763 GATCCCAAGCCTCAGCGGGTGGTGGAGTCCCTTATGTTCTCTCACCA 1812  
 250 nProThrTrpAspCysGlnGlnLeuLeuGlnThrLeuPheThrhrG 267  
 1813 GCCTACTGCGATGATGTCACACACTCTGACAGACTCTTCACACCG 1862  
 267 LgluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
 1863 AGGAGCGAGAGAAATTCGTAGAGGCTAGAAAAAATGTTCTGGGGCC 1912  
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeuLeuLeuLeuLeuLeu 300  
 1913 GAGGGGCGACCGACGAGTTGCAAAATGATGATTCACATGGATTCCCTT 1962  
 300 uThrArgProGlyTrrpAspTrrpAsnThrAlaGluLysArgGluSerLeu 317  
 1963 GACTGCCCCGGTGGAGCTACACACGCGCTGAGGTAGAGGAGACCTTGA 2012  
 317 yAlleYrrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 2013 AAATCTATGCGAGGCTCTGGTGGGGGCTCCGGGGGCGCTTCACAGACGG 2062  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 2063 CCACATAATTGGCTAGCTAGAGAGGTGATGACGAGGAGCCACAGAACCC 2112  
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgPheThrP 367  
 2113 TCCCTCGATTATCTTGGAGAGGCTCATGAGACCTTCAGCGGTTCCACC 2162  
 367 rPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 2163 CTTTATCTTACCTCGAGAGGCCAGAAAGCCTCACTGAGCTTC 2212  
 2213 ATTGGGACAGTGGCTCTGATATCAGAAAGAACTTCAGACACTGAGAGG 2262  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValT 417  
 2263 GTTACAGAGGCTGAGTTCGTATCTAGTACAGAGAGGACAGAAAGTGT 2312  
 417 yTrrArgArgGluThrGluGluGluLysGluGlnArgLysGluLysGlu 433  
 2313 ATTACAGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2362  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 2363 AG 2412  
 450 rLysIleLeuAlaAlaValAlaGluLysSerSerArgGluArgGluAla 467  
 2413 TAGATCTTGGCCGACGTGTTAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 2462  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 2463 GAGATTTTAGGAAATAGGTACAGGCCCTAGACAGTACAGAGAGAGAGAG 2512  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500  
 2513 AATTAGACCCCTCAGCAAGACCACTGTGCGATTGTAAAGAAAAAGG 2562  
 500 yHisTrrAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValL 517  
 2563 ACACCTGGGCAAGAGACTGCCCAAGAGAGAGAGAGAGAGAGAGAGAGGTC 2612

517 euAlaLeuGluGluAspLysAsp 524  
 2613 TAGCTCTAGAGAGAGATTAAGAT 2635

seq\_name: gb\_v1: PEN293657

seq\_documentation\_block: 8763 bp DNA VRL 30-MAY-2001

LOCUS PEN293657 Porcine endogenous type C retrovirus proviral gag gene, pol gene

DEFINITION and env gene, class B, clone 213.

ACCESSION AJ293657.1 GI:14275841

VERSION env gene; envelope; gag gene; group specific antigen; pol gene;

KEYWORDS polymerase; polyprotein.

SOURCE porcine endogenous type C retrovirus.

ORGANISM Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.

REFERENCE 1 (bases 1 to 8763)

AUTHORS Krach, U., Fischer, N., Czauderna, F. and Tonjes, R.R.

TITLE Comparison of replication-competent molecular clones of porcine

JOURNAL endogenous retrovirus class a and class b derived from pig and

human cells

U. Virol. 75 (12), 5465-5472 (2001)

21256017

REFERENCE 2 (bases 1 to 8763)

AUTHORS Tonjes, R.R.

TITLE Direct Submission

Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology,

Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,

GERMANY

#### FEATURES

source

Location/Qualifiers

1..8763

/organism="porcine endogenous type C retrovirus"

/proviral

/specific\_host="Sus scrofa"

/db\_xref="taxon:70540"

/clone="213"

/cell\_line="PK15"

/note="class B"

1..630

467

/note="Cap site"

633..650

1077..2651

/gene="gag"

1077..2651

/gene="gag"

/codon\_start=1

/product="group specific antigen"

/protein\_id="CAC39625.1"

/db\_xref="GI:14275842"

/translation="MGQVYTPPLSLTLDHMTVEYRAHNLVQVKKGPWOTFCASEMP

TFDVGMPSEGTENSEILIAVKAIIFOTGSHDDEPYILTWDLAEEPPWPKPWLN

KPRKPGRIALIEKNKHSAEKVPSPGILYELIEPPWPEQPVPPYPVAGAVRG

PSAPPCAPAVEGPAAGSTRGATPRTEIDALILPRTYGPMPGGLOPLQVWPS

ADLYNKTINRPSPEDPORLTVGESIMSHOPTMDCOLLQTLTTEREELIEA

RKNVPGDGRPTOLDNEDIDMGFPPLTRGMDYNNAEERESIKYIRKALVIGRASRP

TNLAKVEWVGQENEPVSFLERLMEAFREFTFEDTSEQRASVALATIGOSALDIR

KILQREGLOEALRDLYRAEVYVRRTEERKEQREKEREERERERDRROEKVLT

KILAAVEGKSSRERERDFKIRSGPROSGNLGNRTPLDKDCAVCKEKGHWARNCPK

KGNGKPVLAIEDKD"

2652..6239

/gene="pol"

<2652..6239

/gene="pol"

/note="complete pol ORF, gag-pol is expressed as a

polyprotein"

/citation=[1]

/codon\_start=1

/product="polymerase"

/protein\_id="CAC39626.1"

/db\_xref="GI:14275843"

/translation="GRRGSDPLPEPRVTLKVEGQPVFELVDGAEHSVYLQPLGLIKE

```

KGSWVGATGCGQCYQPMTRRTTADLVAGVTHSFLVTPRCPCVPLIGRDLITWMAQGIS
ECRSMVAVNNKPIVVLITLDDDEVRKLSPOVKPRODQISMLDOPRPMQATAGGLA
ROKPPVOVIOIKASATPVSQVROVPLSRKARBEQIMPHVQILIOGGIIVPQSWMTPLIL
VAKPCTGNIDRIPODLREVNKRQDILHTRIPMDYNLSLPERMMWYVLADLKAFECI
RLHPTSQPLFAEWMKDPGCTGKRVQDILHTRIPMDYNLSLPERMMWYVLADLKAFECI
VTLIDYVDLDDLLAGATKQDCLEGTAKALLLESDJGVRASAKAKOICREVEYLGSLSR
GGOURLTEARKKTAVOIPAPTAQKQVREFLEGTAGCRLMIGFPAITLAPILYLRKRGK
EFSMABEHQAKFADJAKKALLSAPALADLEVRKRPPLIYVDEKRGVARGVLTOTIGPMWR
PVAYSKLDIPASGMPYCICAIAVAIVLIVMDACKTLTGQNTIYTAHAIENLYRQPP
DMMNNAHRTHQSLILTERVTFAPRALNATLILPETTDVPIHDDHOLIETGVK
KLOTPIPLGEVLTPTDGSSTVBEKMGACAADVNDGRTIWSARELBEQISQAELM
ALTQLRABEGKSTNIYDTSRAFTAVHVAIYKORGLTASARELBEQISQAELM
ALHLEMRJAILHCPGHOAKDLISKGNOMADRVAKQAQAVNLPIETPRAPRRRO
VTLIEMOEIKRIDQSEPESTGCTSYSGKELLPHKEGLEVYQOILHRLHGTKLQOLO
VTSYHYVLRLEPGVADSVKACVPCOVLNANPSPRGKRLRGSHPGAAHVEDPTEVK
PAKYNSKTLIVFVDFESGWEAVYPRKKTSTVYAKKILIEELFFRFGILPKYIGSDNGFA
FYAOVSQGLAKILGIDMKLHCAYRQSSGOVERNKRITKEVTLTETEGINDIALL
PVLVLRVNTPOQFGLPYELLVGGPPVLEIVASVHSDVLLSQPLSRILKALEMVRQ
RAMKOLREAYSGEDLOVPHRFQVGDSSVYRHHAGNLEIRMKQPSVLLTPTPAVKY
EGISTWIIASHVVKLAPPDSGMRAEKTEINPLKRLHRLVPSXNSNNGQ"

gene
6112..8085
/gene="env"
6112..8085
/gene="env"
/odon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/translation="MHPTLSWRHLPTRGSGPKRLRILPSPASTAMPLTLITTPQASSK
RLIDSNPHRPLSLTLIIPDDCVTYNSTRGVAPRQETWMELEHFCRLINPAVKSTR
PNLVRSTCYCPGCTEKEKCTGGSGESFCRMSKSTNSDNDMKWPLSLDORVAFESVN
SGPKNYAKMLYKDKSCSPDLMDYLKLSFTEKGEQENIOWIMNGMSWGLVYKYGGGA
GSTLILIRLETETEPVAVGPPKVLAEQGPALPEPRLNIPOLNLSRPDIPOPPX
STTOLIPETNPNSGPPVKTQGLRFLIGAGRAIINSTPDALTSQWCLSSGPPY
EGMAKEKENVTKERHNOCTMGSSRNKLTLENGEAGCTGCKAPSHOHCNSVYVEQ
ASENQIVPGVNMWMACTNGLTPCVSTVSNOCKDECVMOIYPRVYIPEEVYLDIY
DYARNRKRREPVSTLAVMLGLTAIVSGTGLATITLPGQDLKGJGSEYHAAMTEDLR
ALLESVSNLEBSLTSLEVVYLDNRGIDLFLREGLGCAALKECCFPYDHSGAINDS
MSKRLERLERRRRENDQGMFEGMFNRSPMTLILALGPLVVLILLITVGPCLIN
RFVAFERERVSAQVIMVLRQOYGLISQGETLD"
8133..8673
8655..8660

LTR
polya_signal
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN

alignment_scores:
Quality: 2752.00 Length: 525
Ratio: 5.262 Gaps: 1
Percent Similarity: 99.619 Percent Identity: 99.238

alignment_block:
US-09-171-553b-4 x PEN293657 ..
Align seg 1/1 to: PEN293657 from: 1 to: 8763

1 MetGlyGlnThrValThrThrProLeuSerLeuThrIleuAspHisThrPth 17
|||||
1077 ATGGGACAGACAGTAGTACCCCTTAGTTGACTTCGACCAATTGGAC 1126
|||||
17 TGIuValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
|||||
1127 TGAAGTTAGATCCAGAGGCTCATAAATTGTGACTTCAGGTTAAGAAGGAC 1176
|||||
34 rotpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
1177 CTGGGCAAGCTTCTGTGCTCTGTAATGGCCMAATTCGAGTGTGATGG 1226
|||||
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaIle 67
|||||
1227 CCATCAGAGGGAGCACTTAAATTCGAATTATCTCGCTGTAAAGCAAAAT 1276
|||||
67 eilepHeGlnThrGlyProGlySerHisProAspGlnGluProTyrIleIle 84

```

1277	CAATTTCAGACTGGACCCGGCTCATCTGTATCGAGGCCCTATATCC	1326
84	euThrTrpGlnAspLeuAlaGluAspProProtrPylLysProtrP	100
1327	TTAGCTGGCAAGATTGTGGCAAGAAACCTCCGGCAATGGGTAAACCAATGG	1376
101	LeuAsnLysProArgLysProGlyProArgLysLeuAlaLeuGlyGlyLys	117
1377	CTAAATTAACCAAGAAAGCCAGGTCGCCGATCTGGCTTGGAGAGAA	1426
117	SAsnLysHisSerLysAlaLysValGluProSerSerSerLysProA	134
1427	AAACAAACACTCGGCCAATAAAATCGACGCCCTCTCTCTGG..ATATCAACCC	1475
134	rgAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProthr	150
1476	GAGATCGAGGAGCCGCCGACTTGGCCGGAAACCCAACTGTCCCCCACC	1525
151	ProLysSerSerThrGlyCysGlyLysThr..SerAlaProGlyA	167
1526	CCCTTATTCAGACACAGGGTCTGTGGAGGACCTCTGCCCTCTCTGGAG	1575
167	laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla	183
1576	CTCCGGTGGGAGGACCTGTGGCCGGGACACGGAGCCGAGAGGGGCC	1625
184	ThrProLysArgThrAspGlyLysAlaAlaLeuProLysAlaGlyTrpGly	200
1626	ACCCGGAGGGGACACACGAGATCCGAGATTACCGCTGGCCACTTATGG	1675
200	yrProPrometPProGlyGlyGlnLeuGlnProLeuGlnIleTrpProbes	217
1676	CCCCCCCATCCCAAGGGGGCAATTGACGCCCTCCAGATTATGGCCCTTTT	1725
217	erSerAlaAspLeuLysTrpAsnTrpLysTrpAsnHisProProPhesSerGlu	233
1726	CTTCTGCAAGATCTCTTAATTGGAAACTAAACATCCCCCTTGTCCGAG	1775
234	AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly	250
1776	GATCCCCCAACGCCCTCAGGGGTTGGTGGAGTCCCTTAATGTTCTCTACCA	1825
250	nrProThrTrpAspAspCysGlnGlnLeuGlnGlnThrLeuPheThrThrG	267
1826	GCCTACTGGGATGTGTCAACACCTCTCAACAACCTCTTCAACAACCG	1875
267	LuGlnArgGluArgLysLeuLeuGlnAlaArgLysAsnValProGlyAla	283
1876	AGGAGGAGAGAAATTGCTTGAAGGCTTAACAAAAATGTTCTCGGGGCC	1925
284	AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle	300
1926	GAGGGGCGACCCACAGTGTGCAAAATGAGATTGACATGGAGATTTCCTT	1975
300	uTrpArgProGlyTrpPAspPyrAsnThrAlaGluArgLysSerLeuL	317
1976	GATCGCCCCCGGTTGGGACTCAACACGGCTGAAGGTAAGGAAAGCTTAA	2025
317	ysLleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg	333
2026	AAATCATATGCGACAGGCTGTGGTCGGAGTCCCGGGCGGCTTCACAGCG	2075
334	ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr	350
2076	CCACATTAATTGGCTAAAGTAAGAGAGGTATGACAGGACCCGAACCAACC	2125
350	oerOserValPheLeuGlnArgLeuMetGlyAlaPheArgArgPheThrP	367
2126	TCCCTCGGATTCTTGAAGGCTCATGAAAGCCTTCAGGCGGTTCCACC	2175
367	roPhaAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe	383

```

2176 CTTTGATCTACCTCAGAGGCCGAGAAAGCCTCAGTGGCCCTGGCCTTC 2225
384 IleglyglnSerAlaIleuAspIleArgIlyLeuGlnArgIleuGlu 400
2226 ATTGGGCGATGGCTGCTGATATCAGAAAGAACTTCAGAGCTGAGAGG 2275
400 yLeuGlnIleuAlaIleuArgAspLeuValArgIleuAlaGluIlyValIT 417
2276 GTTACAGGAGGCTGAGTTACGATCTAGTGAAGAGGACAGAGAGGAGTGT 2325
417 yTyrArgArgGluThrGluGluGluGluGluGluGluGluGluGlu 433
2326 ATTACAGAAAGGAGACAGAGAGGAGAAAGAAACAGAAAGAAAGAGAG 2375
434 ArgGluGluArgGluGluArgAspArgArgGluGluGluGluGluGlu 450
2376 AGAGAAAGAAAGGAGAGAAAGAGCTAGACGGCAAGAGAAAGAAATTCAC 2425
450 IlyIleIleuAlaIleuAlaIleuGluGluGluGluGluGluGluGlu 467
2426 TTAGATCTTGGCCGAGTGTGTAAGGAGAGAGAGAGAGAGAGAGAGAGA 2475
467 rGAspPheArgIlyIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
2476 GAGATTTTGAAGAAATTAGCTCAGGCCCTAGACAGTCAAGGAACTGGGC 2525
484 AsnArgIlyProLeuAspIlyAsnGlySalIaIyrcIlyGluGluGlu 500
2526 AATAGACCCCACTCGACAAAGAGACAGTGTGCTATTGTAAGAAAGAAAG 2575
500 yHisTAlaArgAsnCysProIlyGlyGluGluGluGluGluGluGluGlu 517
2576 ACACCTGGCAGAGAACTGCCCAAGAGAGAAACAAAGAGAGAGAGAGTCC 2625
517 euAlaLeuGluGluAspIlyAsp 524
2626 TACCTTAGAGAAAGATAAAG 2648

seq_name: gb_v1:PEN133816

seq_documentation_block:
LOCUS      PEN133816      8918 bp      DNA      07-DEC-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 33).
ACCESSION AJ133816
VERSION    AJ133816.1 GI:6688945
KEYWORDS   env gene; env protein; gag gene; gag protein; pol gene; pol
           protein.
SOURCE      porcine endogenous retrovirus.
ORGANISM   Porcine endogenous retrovirus
            Viruses; Retroid viruses; Retroviridae; Mammalian type C
            retroviruses; 1-Mammalian type C virus group.
REFERENCE   1 (bases 1 to 8918)
AUTHORS     Czadudena,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE       Establishment and characterization of molecular clones of porcine
            Establishment and characterization of molecular clones of porcine
            endogenous retroviruses replicating on human cells
JOURNAL     J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE     20219394
REFERENCE   2 (bases 1 to 8918)
AUTHORS     Toenjes,R.R.
TITLE       Direct Submission
JOURNAL     Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
            Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
            Langen, GERMANY

FEATURES
            source
            1..8918
            /organism="porcine endogenous retrovirus"
            /proviral
            /strain="type C"
            /db_xref="taxon:61673"
            /clone="33"
            /note="grown in PK15 (ATCC No. CCL-33) cell line"
            1..707

LTR

```

---

```

545
/note="cap site"
710..727
/note="91y4"
1154..725
/note="gag"
1154..725
/note="gag"
/function="retroviral capsid"
/product="gag protein"
join(1154..725,729..6313,6189..8162)
/note="viral polypeptide"
/codon_start=1
/protein_id="CAB65339.1"
/db_xref="GI:6688946"
/db_xref="SPTREMBL:O901X5"
/translation="WGQVTVTTLSTLDHWTEVRKRNLSVQKKGGPOTFCSEWP
TPDVGPSSEGTENSEILAVKALIFQTPSPSHDDEPYLTLMODLADPPVAKFWLN
KRRKGPRLALGKRNKSAEKVPSRPRIPEIEPPWPEPOVPVPPVPAQAVRG
PSAPGAPVVEGPAAGTAKSRGATPERIDEALILRLTYGPPMGGQLQLOYPFSS
ADLYNMKTNHPSPBPORLRLGVSLSHESHPQWDCQOLLOTFTTEBERLILEA
RKNVGCADGRPTOLONEIDMGPRLRQMDVNTAGRESIKIYRALVAGRGASRP
TLNAYREVWQGNPSPVTEFLRMEARRRTTPPTSEAKASVALPFGSALDIR
KLIQLEGLQELADLVNKAERYRRETEKQREKREBERERDRROKCNLT
KILAAVECKSSRERDRERKIRSGPDSGNLGNRPDLKDCACVCKRGHAAKCPK
KGNKGKVLILEEDDGRGSDPLPEPRVTLKVEQVPEFLVDGAEHSVLQPLGLK
KKSQWVGATGOROYPTTRTVDLAVGTHSELVIPCEVPLIGDLTKMGAQI
SPGQREYSNNKPTVTLTLODEYRLSPQVDDOIOSMLBOFOAMETAGMG
LAKOFPVOYIOIKASATPVSAQVYPLSRAREGIAPHORLIOGLIVPVSPPMTPL
LVRRKPGTNDYRPVDLREVRKRVODIIFTVNPYNLSALPEERNWTVDLKAPF
CLRLHETSOPLPAFMRDGRTGRTQLWTRLPQCFKNSPTLFEDALHRLANFRIQ
PQVTLIOYVDDLLAGATKODLEGTKALIELSOLGRASAKAOICORREVTLYGS
LGGORWILEARKTYVOIPAPTTAKOVEEFGTACFRMLPGFATLAPLYLTKE
KGEFSAPHEOKAPDAIKKALISAPALPDVTKPEPTLYDERKAVAGVLTQGLPW
RRPVATLSKLDIPVASGWPVCUKATAAVALYKDDAKLTGNTIVIAPHALENTVRO
PPDRMNTNKMTHYOSLLTERVTTPAPALNPATLLEPTEDEPTTHOCHLIEETG
VKRDLTDIPLTGEVLTWFTDSSVYVEGRHAGAAVVDGTHIMASSLPEBGSARAE
LMALTQALRILABGKSLNITYDRAVAFATVAHAGIYKOGSLTASGRETKNEELSL
LEALHPRLATIHCHGHOKRADLISRGNOAMDRAVAKQAQVNLPIETPKAREPR
ROYTLEDMOELKKIDOFSETPEGTCYTSYKELILPKBELEVVOGILHRLHGTAKHLQ
OLVTPSPYHVLRLPGVADSVKHCWCOLVYNNPRLIPGKRLRSGSHGAMHEVDTE
VPAKYGKXYLLAVYDTPESGWEATPTKRETSYVAKKILEETPRFGIPKYGSDNG
PAPVAVOSGLAKILGIDMKLHCARPOSSGOVERNRTIKETLTKLTETGINDMIA
LLEFVLEVRNTPPGGLTPRYELLYGGPPLEIASVSHADVLISQPLESRKALEWV
KORAMKOLREAYSGEDLOVPRFOVGDSVYVRHAGNLETRMGVYLVLTTPFAY
KVEGISTWITRSHVLAAPPDGSRAEKTENPLKRLHLVPSYNNKNSGQIHPTLSW
RHLPTRGGEPRKRLRPLSPASTAMFELTLTPQASSKRLIDSSNHRPLSLTWLIIDP
DTGVYVNSTRGVAPRGCTWMPLEHLCRLINLPVAVKSRPMLVRSYGCOPGKEXKC
GDSGESFCRMSCVTSNDGDKWPIISLQDRVFSYVNSGPKYKWKLYKDKSCSPSD
LDYLKISFTEKQENIOKINGMSWGLVIFYXGGAGSTLIRLLETGTPPAVAG
PDKVLAEGPPALEPHNLPPQLSLRDLIOQSPNGTGLIPNTPRNSGVPVKT
GORFLSIGAFOAINSTDPAATSCMLCISGPPYEEMAEKGFENYKREKNCOTW
GSRNKLITLIEVSGKTCIGKAPSHOHLGSIYVYVYQASAEQYLVAGYRRMACTGL
TPCVSTSVNOSKDRCMVOLEKGLGLAAMTEDIRALEEVSNTSEBTSLSAVGL
ONRAGVIGTALITLTPQOLEKGLGLAAMTEDIRALEEVSNTSEBTSLSAVGL
LGNRAGDILLFLRGGICAAKKECCFYVHSGAIRSMKLERLRRRREBADOGW
FEQGLSFRSPPMTTLISALTGLPVLVLLLTVPCLINRPVAVREKRSVAVOIMVLRQ
YQGLSFRSPPMTTLISALTGLPVLVLLLTVPCLINRPVAVREKRSVAVOIMVLRQ

```

---

```

misc_feature
primer_bind
gene
mat_peptide
CDS

```

---

```

mat_peptide
gene
mat_peptide
gene
LTR
polyA_signal
BASE COUNT      2327 a      2245 c      2334 g      2012 t

```

## ORIGIN

## alignment\_scores:

Quality: 2751.00 Length: 525  
 Ratio: 5.250 Gaps: 1  
 Percent Similarity: 99.810 Percent Identity: 99.429

## alignment\_block:

US-09-171-553b-4 x PEN133816 ..

Align seg 1/1 to: PEN133816 from: 1 to: 8918

```

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh
1154 ATGGGACAGACAGTGCAGACCCCTTACTTAACCTCGACCAATTGGAC
1203
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP
1204 TGAAGTTAAGTCCAGGGCTCATTAATTGTTCAGTTTCAGTTAAGAGGGAC
1253
34 rOTrPGLnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp
1254 CTGGCAGACTTCTGTGCTCTGATGCAATGCCAACATTCGATGTGGATGG
1303
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI
1304 CCATCGAGAGGGACCTTAAATCTGAAATTAATTCCTGGCTTAAAGCAAT
1353
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL
1354 CATTTTTCAGACTGGACCCAGCTCTCATCTCATGAGAGCCCTATATTC
1403
84 euThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp
1404 TTACGGCGCAGACACTTGGCAGAAATCCCGCCCATGGGTTAAACCATGG
1453
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLysL
1454 CTAAATAAACCAAGAAAGCCAGGTCGCCGAAATCCGGCTTGGAGAGAA
1503
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA
1504 AAACAACACTGGCGGCAAAAAGTGGAGCCCTC.TCCTCGATCTACACCC
1552
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr
1553 GAGATCGAGGAGCCCGCACTTGGCGGGAACCCCAACCTGTTCCCCAC
1602
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGlyA
1603 CCTTATCCAGACACAGGGTCTGTGAGGGGACCCCTGCTCCCTCTGGAG
1652
167 laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAlaA
1653 CTCGGGTGGTGGAGGACCTGCTGCGGGGACTCGGAGCGGAGAGGGCC
1702
184 ThrProGluArgThrAspGluIleAlaAlaLeuProLeuAspGlyTrpG
1703 ACCCGAGGCGGACAGAGAGATGCGCATATTACCGCTGCGCACCTATGG
1752
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrCThrProPhes
1753 CCCTCCCATTCGGCGGGGGGCCAATTGCCAGCCCTCCAGATATGGGCCCTT
1802
217 erSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPhSerGlu
1803 CTCTCGAGATCTCTATAATGGAACACTAACCATCCCTTCTCTGGAG
1852
234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG
1853 GATCCCAACGCTTCACGGGGTGGTGGAGTCCCTATGTCTCTACCA
1902

```

```

250 nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrTrg
1903 GCCTACTTGGGATGATGTGCAACACCTCTGACAGACACTTTCACACCG
1952
267 luGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla
1953 AGGACGAGAGAGAAATCTGTAGAGGCTAGAAAAAATGTTCTGGGGCC
2002
284 AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLe
2003 GACGGGCGACCCACGACAGTTCGCAAAATGATGATGACATGAGGATTTCC
2052
300 uThrArgProGlyTrpAspTyrAsnThrAlaGluIleValArgGluSerLeu
2053 GACTGCGCCCGGTGGTGGACCTACACACGCGCTGAAAGTATGGAGACTTGA
2102
317 ySleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg
2103 AAATCTATCGCCAGGCTCTGTGGCGGCTCCGGGGCGCCTCAAGACGG
2152
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluTr
2153 CCCTAATTTTGGCTAAAGTAAAGAGATGATGACAGGACCGAAGCAACC
2202
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP
2203 TCCCTCGGTATTTCTTGAAGAGCTCATGGAAGCTTACGGCGGTTCACCC
2252
367 rOPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe
2253 CTTTGAATCTCACTCGAGAGCCCGAAGACCTCAGTGGCCCTGGCCTTC
2302
384 lIleGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluG
2303 ATTGGGAGCTCGGCTCTGATATCAGAAAGAACTTACGAGCTGGAAGG
2352
400 yLeuGlnGluAlaGluLeuAspLeuValArgGluAlaGluLysValTr
2353 GTTACAGAGAGCTGAGTTAGCTGATCTAGTGAAGAGCGCAGAAAGTGT
2402
417 yTyrTrpArgArgGluThrGluGluGluLysGluGlnArgLysGluLysG
2403 ATTACAAAGGAGACACAGAGAGAGAGAAAGCAAGAAAAAGAGAG
2452
434 ArgGluGluArgGluArgGluArgGluArgGluGluLysAsnLeuTh
2453 AGAGAAAGAAAGGAGAAACCTGATAGACGCAAGACAGAAATTTGAC
2502
450 rLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGluA
2503 TAAGATCTTGGCCGAGTGGTTGAAGGAAAGACACAGAGAGAGAGA
2552
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlnAsnLeuGly
2553 GAATTTTAAAGAAATTAAGTCAAGCCCTAGACAGTCAAGGAACTGGGCG
2602
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG
2603 AATAGAGCCCACTGACAAAGACCAAGTGTGCTATATTGAAGAAAAAG
2652
500 yHisTrpAlaArgAsnCysProLysLysGlyLysAsnGlyProLysValL
2653 ACACCTGAGCAAGGAACTGCCCAAGAGGAAACAAAGAGACCAAGGTCC
2702
517 euAlaLeuGluGluAspLysAsp
2703 TAACTCTAGAGAGATTAAGAT
2725

```

seq\_name: gb\_v1:PERY17013

seq\_documentation\_block: 7808 bp DNA

LOCUS PERY17013 VRL 14-APR-2000  
 DEFINITION Porcine endogenous retrovirus gag, pol, env genes.





809 ACCCGGAGCGGACAGACGATCGGATATACCGCTGGCCACTATGG 858  
 200 YPRPROMETPROGLYGLYINLEUINPROMLEUINLYRTPPROPHES 217  
 859 CCTCCCATATCCGGGGGGCAATTGGACGCCCTCCAGTATGGCCCTTT 908  
 217 eRserAlaAspLeuTYRAsnTPRPLYSThrAsnHisProProPhesArg 233  
 909 CTTCTGCAGATCTCTATTAATTTGAAAACTAACCATCCCTCTTCGGAG 958  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250  
 959 GATCCCAACAGCCCTCACGGGGTGGAGTCCCTATATGCTCTCCACCA 1008  
 250 nPRoThrTPAspAspCysGlnGlnLeuLeuInThrLeuPheThrThg 267  
 1009 GCCCTACTTGGAGATTGTCAACAGCTGCTGCAGACACTCTTCACAGCCG 1058  
 267 LUGLUAArgGluArgGlyLeuLeuGluAlaArgLYsAsnValProGlyAla 283  
 1059 AGAGCGAGAGAGAATTTCTGTAGAGCTAGAAAAAATGTTCTGGGGCC 1108  
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeasPmetGlyPheProle 300  
 1109 GACGGGCGACCCAGCAGTTGCCAAATGAGATTGACATGGGATTCCTT 1158  
 300 uThrArgProGlyTrpAspTYRAsnThralGluGlyArgGluSerLeu 317  
 1159 GACTCGCCTCGGTTGGACTACACACGCGGTGAAGTAGGAGAGCTTGA 1208  
 317 ySIEtyrArgGlnAlaLeuValaGlyLeuArgGlyAlaSerArgArg 333  
 1209 AATATATCGCCAGGCTCTGGGGGGCTCCGGGGGGCTCAAGAGG 1258  
 334 ProThraNleuAlaLysValaArgGluValMetGlnGlyProAsnGluPr 350  
 1259 CCCACTAATTTGGCTAAGGTAAAGAGGTGATGCAGGACCGCAAGAAC 1308  
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgRphrHP 367  
 1309 TCCCTGGTATTTCTTGAGAGGCTCATAGAGCCCTTCAGGGGGTCAACC 1358  
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValaLalaLeuAlaPhe 383  
 1359 CTTTGTATCTTACCTCGAGAGCCAGAAAGCTCAGTGGCCCTGCTTC 1408  
 384 ILeglyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400  
 1409 ATTGGGAGTGGCTGTGATATCAGAAAGAACTTCAGAGACTCGAAG 1458  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417  
 1459 GTTACAGAGAGCTGAGTTAGCTGATCTAGTGAAGAGGAGAAAGGCT 1508  
 417 yTYrArgArgGluThrGluGluLysGluGlnArgGlyGluLysGlu 433  
 1509 ATTAAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558  
 434 ArgGluGluArgGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 1559 AGAGAAAG 1608  
 450 rLYsIleLeuAlaAlaValaGluGlyLysSerSerArgGluArgGlu 467  
 1609 TAGAGTCTGGCCGAGTGGTTAAGGAGAGAGAGAGAGAGAGAGAGAG 1658  
 467 rGAspPheArgGlyIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 1659 GAGATTTTGAAGAAATAGGTACGCCCTAGACAGTCAAGAGAACTGGGC 1708  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTYrCysLysGluLysG 500  
 1709 AATAGAGCCCACTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758

500 yHISrPAlaArgAsnCysProLysGlyAsnLysGlyProLysVal 517  
 1759 ACACGGGCGAAGAGACTCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1808  
 517 euAlaLeuGluGluAspLysAsp 524  
 1809 TAGCTCTGAGAGAGATTAAGAT 1831

seq\_name: gb\_v1:PEN293656

seq\_documentation\_block:

LOCUS PEN293656 8918 bp DNA VRL 30-MAY-2001  
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene  
 and env gene, class A, clone 58.

ACCESSION AJ293656  
 VERSION AJ293656.1 GI:14275837

KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;  
 polymerase; polypeptide.  
 SOURCE Porcine endogenous type C retrovirus.

ORGANISM Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.

REFERENCE 1 (bases 1 to 8918)  
 Krach, U., Fischer, N., Czauderna, F. and Tonjes, R.R.  
 Comparison of replication-competent molecular clones of porcine  
 endogenous retrovirus class a and class b derived from pig and  
 human cells

AUTHORS J. Virol. 75 (12), 5465-5472 (2001)  
 MEDLINE 21256017

JOURNAL 2 (bases 1 to 8918)  
 TITLE Direct Submission  
 AUTHORS Tonjes, R.R.  
 SUBMITTED (31-Jul-2000) Tonjes R.R., Medical Biotechnology,  
 Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,  
 GERMANY

FEATURES  
 source

Location/Qualifiers  
 1..8918

/organism="porcine endogenous type C retrovirus"  
 /proviral  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:70540"

/clone="58"  
 /cell\_line="PK15"  
 /note="class A"

1..702  
 467  
 /note="Cap site"

706..723  
 1153..2727  
 /gene="gag"

1153..2727  
 /gene="gag"

/codon\_start=1  
 /product="group specific antigen"

/protein\_id="CAC39616.1"

/db\_xref="GI:14275838"

/translation="MGQVYTPPLSLFDHMTVEKSRANLSYQKKGPWQTCSEMP  
 TFDVGPSECTFNSLILAVKATIPQTGSHPDQPYLTQDLADEPPTVWTKFWLN  
 KPRPGRIALGEMKHSAEKRPSPHYIPELEPPAPPEQSVPPPYPAQAGAR  
 PSAAPRAPEVGAAGTSGRGTPEPTEIATLPLRYGPPIPGGQLDLETTERKILLEA  
 ADLYNNKTNHPESDPPORLTGLVESLMFSHQPTWDCOQLQTLTTEERKILLEA  
 RNKVPADAPPTOLONEIDMGPELTPRGMDYVTAERESLKYRQALVGLGASRRP  
 TNLAKVREVMQGNPEPSVLELMFAFRFTPPDPTSAOKASMAIAFICGSALDIR  
 KILQREGLQEAELHDLVKEAEKVVYKRETEERERORREBERBERNNKQENLT  
 KILAAVEKSNRERDRPKIRSGRSGNIGNRNPIDKDCAYCKEKGHARDCPK  
 KGNKGILKVALBEDKD"

2728..6309  
 /gene="pol"

<2728..6309  
 /gene="pol"

/note="complete pol ORF, gag-pol is expressed as a  
 polypeptide"

/codon\_start=1

gene

CDS

gene

CDS

gene

CDS

gene

```
/product="polymerase"  
/protein_id="CAC39617.1"  
/db_xref="GI:14275839"  
/translation="GRGSDPLPEPRVTLKEGQPEFLVDGAKHSLLDPLKGLKD  
KSSWVGATGQGOQPMTRRVDLVGVGRVTHSFLVPECPALPLGRLLDKMAQISF  
EOKGPEVSANNKPTVTTLQDDEYRLXPLVPRDQIQWILQFOPMAEETAGMGLA  
KQPPQVYQLKASAPVSVKQPLSKEAREGIRHVRQLIQQGLVLPVQSPNTPILP  
VRKPNIDYREVQDIREVNRVODIHPTVPNPVNLCLAPQKRWTVLTLKAFQCL  
RLHPTSQLPEFWRDPCAGRTGQTLWTRLPOGKNRPTLFEDLAHDLNFRIGIQRH  
VTLLQYVDLLLAGATKQDCEGTAKALLLESDYASAKKQICRETEYVIGYSIR  
GGORLIEARKRRVVOIPPTAKOVRREFLTGFCGLMIGFATLAAPLYLPIKKG  
EESWAPHEOKAFDAIKKALLSAPALAPDYTKPTLYLVDKRYKARVLTQIGPBR  
PVATLSKLDPVASGWPICLKAIAVAALVADKLTGQNTVIAVHALENTVPRP  
DRMTNARMTHYOSLLTERVTFAPRALNAPLLEPEIDEPYTHDQDLLEETGVR  
KDLTDILTEGMLTFTDGSYMEGKRMAGAAVDDRTIMASSIPEGISQAELM  
ALPOLRLAEKSIINYDSYAFATVHGAITYKQGLTASGRELKNKEETLSLE  
ALHPKRLATIHCGHOKADPISRGNOMADRAKQACQVNLPMTEHPKAPERO  
YTLDEWQPIKIDPFSTPECTCTSDGKETLPHKEGLETVQQLHRLTLRHLQOL  
VRSIPYHVLRLPGVADSVNRKCVPCQLVNPSPILPEGRKRGSHPGAHMEVPTEK  
PAKYGNKYLLVFDTFSGWVEAYPTKKESTVYAKLLEBIFPFGIPKYGSDNGPA  
FVAVQSOGLAKILGIDMKLCAYRPOSSGOVERMNTIKETLTKLETETINDMMLL  
PEVLEFRNTPGOGFLPYELLGGPPLAEIAFAHSADVLLSOPLESRLKALEWRO  
RAWKOLREAYSGDLOYHREFOVDSYVRHAGNLEPWRKGPVLYLLTTPAVVE  
GITPTWASHVKRPPPDSCMKAEKTEBNPLKRLHVRVPTSVNNSSS"  
6185. .8149  
/gene="eny"  
6185. .8149  
/gene="eny"  
/codon_start=1  
/product="envelope"  
/protein_id="CAC39618.1"  
/db_xref="GI:14275840"  
/translation="MHPILSRHLPIRGKPKRLIPLSEASIAMFLTSLTPOVNGK  
RLVDSPSNHSKPLSLTSLTSGTGINSTGKPELGTWPELYVCLRSVTPGLNDQA  
TPPDVLRAYGVYCPGPNNEEYCGNODPFCOMSTITSDNCKMPPVSCODRVYS  
EVNPTSTNPNNGHGRKMDQKQVQKDVNNKQISCSLSDLDIKLSFTEKQENIQ  
KWVNGISWGLVYGGSGRKKSVTLRLIETEMEPVALGPKNGLAEGQPPLOEORP  
SPNSDNTNYSVPTENITIKTGAKFLIOGAFALNSTPEKATSSCMALASGP  
PYEGMARGKFNVTKEHRDQCTWGSQNKLLLEVGKSGICIMVPPSHQILCNHTEA  
FNRTSESOVLVPGDWMACNTGLTPCVSTLVNQTDFCOMOIVPRYVYEEKVL  
DEVYRNRPKRPISTILAVMLGSLGAAVGTALTALTPQOOLKESLNRIYTE  
DLQALEKSVSNLESLTSLSEVYLQNRGLDLFLKGGICVALLKECCCTVYHSGAI  
RDSMSKLRELERKHEKEAGQGFEGFMRKSPWYTTLLSALGPLVILLTLTVCB  
LINFVAFVREEOVSARINWLRQOYQGLPS"  
8216. .8918  
LTR  
polyA_signal  
8810. .8815  
BASE COUNT 2435 a 2174 c 2253 g 2056 t  
ORIGIN  
  
alignment_scores:  
Quality: 2661.00 Length: 525  
Ratio: 5.127 Gaps: 1  
Percent Similarity: 98.857 Percent Identity: 95.619  
  
alignment_block:  
US-09-171-553b-4 x PEN293656 ..  
Align seg 1/1 to: PEN293656 from: 1 to: 8918  
  
1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17  
|||||  
1153 ATGGGACAGACGGGTGACGACCCCTTATTGACTCTGACCATTCAGC 1202  
17 rgiValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34  
|||||  
1203 TGAAGTTAAATCCAGGGCTCATTAATTTGTACAGTTCAGGTTAAGAAGGAC 1252  
34 rotTpgLntHrPhcCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
|||||  
1253 CTGGGACAGACTTCTGTCTCTGATGATGCGCAGCATTCATGTTGGATGG 1302  
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
```

```
|||||  
1303 CCATCAGAGGGGACCTTAAATTCGAGATTATCCGTGCTGTTAAAGCAAT 1352  
67 eLlPheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84  
|||||  
1353 TATTTTACAGCTGAGCCCGGCTCTCATCCGATCCAGAGCCCTATATCC 1402  
84 eutHrTrpGlnAspLeuAlaGlnAspProProProTyrPValLysProTrp 100  
|||||  
1403 TTACGTGCAGAGATTGGCAGAGATCCTCCGATGGGTTAAACCTTGG 1452  
101 LeuAsnLysProArgLysProGlyProArgGlnLeuAlaLeuGlyGly 117  
|||||  
1453 CTGATTAAGCCCAAGAAACCCAGGTCGCCGAATTCGTGCTTGGAGAGAA 1502  
117 sAsnLysHisSerAlaGlnLysValGluProSerSerTyrLeuPro 134  
|||||  
1503 AAACAAACCTCGGCTGTAAGAAAGTCAGCCCTC.TCCTCATATCTTACCC 1551  
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
|||||  
1552 GAGATTTCAGAGCCCGCGCTTGGCGGGAACCCCAATCTGTTCGCCACCC 1601  
151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167  
|||||  
1602 CCTTATCCGCGACAGGCGTCTGCGAGGGGACCCCTGCGCCCTCTGGAG 1651  
167 lAProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
|||||  
1652 CTCGGGCGGTGAGAGGACCTGCTGCGAGGACTCGAGCGGAGGAGGCGCC 1701  
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTrpG 200  
|||||  
1702 ACCCGGAGCGGACAGACAGATCGCGCATTCGCGTGGCCACGTCAGG 1751  
200 YProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217  
|||||  
1752 CCCCTCCCAACCGGGGGGCCAATTGCAACCCCTCAGATATGGGCCCTTT 1801  
217 eSerAlaAspLeuTyrAsnTrpLysTrpAsnHisProProPheSerGlu 233  
|||||  
1802 CTTCGCGAGATCTCTATATTGGAACATAACCATTCCTCCCTTCTCGGAG 1851  
234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG 250  
|||||  
1852 GATCCCCAGACGCTCAGCGGCTGTGGTAGTCCTTATGTTCTCTACACA 1901  
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThr 267  
|||||  
1902 GCCTACTTGGGATGATTGTCAACAGCTGCTGCACAGACTCTTCAACAACG 1951  
267 lGluArgGluArgIleLeuLeuGlnAlaArgGlyAsnValProGlyAla 283  
|||||  
1952 AGGAGCGAAGAGAAATTTGTTAGAGGCTTAAGAAAATGTTCTCGGGACC 2001  
284 AspGlyArgProThrGlnLeuGlnAsnGlnIleLeuAspMetGlyPhePro 300  
|||||  
2002 GACGGCGACCCACGACGATTCGCAAAATGAGATTGACATGGGATTTCCCT 2051  
300 uThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGlnSerLeu 317  
|||||  
2052 GACTCGCCCGCTGGGGACTGTAACACGCGTGAAGGATGAGGAGGCTTGA 2101  
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
|||||  
2102 AAATCTATTCGCGAGGCTGTGCGCGGCTCTCCGGGCGCTTCAAGACGG 2151  
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
|||||  
2152 CCCACTACTCTGGCTTAAGTAAAGAGAGTGTATGACAGGACGAGATGAAC 2201  
350 oProSerValPheLeuGlnuArgLeuMetGluAlaPheArgArgPheThr 367  
|||||
```

```

2202 TCCCTAGATTTTCTTGAGAGGCTCAWGAAGCCTTCAGGCGGCTCACCC 2251
367 ropheaspProthrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
2252 CTTTATCTCTACCTCGAGGCTCAGAAAGCCTCATGCGCTCGGCTTC 2301
384 ILGLYGLInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
2302 ATAGGACAGTCAGCCCTGATATCAGAAAGAGCTTCAGAGACTGGAAG 2351
400 YLeuGlnGluAlaGlnLeuArgAspLeuValArgGluAlaGlnLysVal 417
2352 GTTACAGAGGCTGAGCTACATGATCTAGTACAGAGGAGCAGAGAAAGT 2401
417 YTYrArgArgGluThrGlnGluGlnLysGlnArgLysGlnLysGln 433
2402 ATTACAAAGGAGGACAGAAAGAAAGGAGCAACAAAGAAAGAGAGAA 2451
434 ArgGlnGluArgGlnLysArgArgArgArgGlnGlnLysAsnLeuTh 450
2452 AGAGAGAAAGGAGGAGAAAGACTAATAACGCAAGAGAGAAATTTGAC 2501
450 rLysIleLeuAlaIleValAlaGlnLysSerSerArgGluArgGlu 467
2502 TTAGATCTTGCTGCTGAGGTGAAGGAAAGCAATAGGAGAAAGAGAA 2551
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlnLysLeuGly 483
2552 GAGATTTTGAAGAAATTAGTGTCAGGCCCTAGACAGTCAGGAAACCTGG 2601
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlnLys 500
2602 AATAGAGACCCACTCGACAAAGCACATGTGATATTGTAAGAAAAAG 2651
500 YHISTPAlaArgAsnCysProLysGlnLysGlnLysGlnLysVal 517
2652 ACAGCTGGGCAAGGAGCTGCCCAAGAGGAAAGCAAGAGCTGAGGTCT 2701
517 euAlaLeuGlnGluAspLysAsp 524
2702 TAGCTCTGAGAGAGATTAAGAC 2724

seq_name: gb_pat:AX052636

seq_documentation_block:
LOCUS AX052636 6076 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 21 from Patent WO0071726.
ACCESSION AX052636
VERSION AX052636.1 GI:12226826
KEYWORDS
SOURCE
ORGANISM
. porcine endogenous retrovirus.
porcine endogenous retrovirus
Retroviruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 6076)
AUTHORS
Federispiet,M.J.
TITLE
Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL
Patent: WO 0071726-A 21 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
source
1..6076
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1613 a 1512 c 1609 g 1342 t
ORIGIN

```

```

alignment_scores:
Quality: 2659.00 Length: 525
Ratio: 5.143 Gaps: 1
Percent Similarity: 98.476 Percent Identity: 95.619

```

```

alignment_block:
US-09-171-553B-4 x AX052636
Align seg 1/1 to: AX052636 from: 1 to: 6076

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17
1183 ATGGACAGAGGCTGACACCCCTTATGTTGACTCTCGACCATTTGGAC 1232
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGly 34
1233 TGAAGTTAATCCAGGCTCATTAATTGTCTCAGTTGAGGTTAGAGGAC 1282
34 rOTrPGLnThrPheCysAlaSerGluThrProThrPheAspValGlyTrp 50
1283 CTGGCAACACTTCTGCTCTGATAGGCCGACATTCGATTCGATTCGAT 1332
51 ProSerGlnGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
1333 CCATCAGAGGGGACCTTAAATTCTGAGATTAATCCTGCTTAAAGCAAT 1382
67 eLLePheGlnThrGlyProGlySerHisProAspGlnGluProThrIle 84
1383 TATTTTTCAGACTGAGCCGCTCTCATCCCAATCAGAGAGCCATATACC 1432
84 eUTrIPGLnAspLeuAlaGlnAspProProPOTrIPValLysProTrp 100
1433 TTACGTGCAAGATTTGGCAGAGGATCTCCGCATGAGGTTAAACCTGG 1482
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlnGly 117
1483 CTGAATTAAGCCCAAGAAACCCAGGTCCCGAATTCGCTCTTGGAGAGAA 1532
117 sAsnLysHisSerAlaGlnLysValGlnProSerSerSerTyrLeuPro 134
1533 AAACAAACACTGCGCTGAANAAGTCAAGCCCTC.TCCTCATATCTACCC 1581
134 rGAspArgGlyAlaAlaAspLeuAlaGlnThrProThrCysSerProThr 150
1582 GAGATTTGAGAGCCGCGCGCTTGGCCGGAACCCCAATCTGTCCGCCAC 1631
151 ProLeuSerSerThrGlyCysCysGlnGlyThr.SerAlaProGlu 167
1632 CCTTATCTGGACACAGGCTCTCGAGGAGGACCTCTGCCCTCTGGAG 1681
167 lAProValGlnGlyProAlaAlaGlnThrArgSerArgArgGlyAla 183
1682 CTCGGCGGTGAGAGGACTCTCTCAGGAGCTCGAGCCGAGGGGCGCC 1731
184 ThrProGluArgThrAspGluIleAlaAlaLeuProLeuArgThrTyr 200
1732 ACCCGGAGCGGACAGAGAGATCGCAGATTAACCGCTGGCCAGATACG 1781
200 yProPROMeTrpGlnGlnLeuGlnProLeuGlnTyrTrpProPhe 217
1782 CCTCCACACCGCGGGGCCAATTCAGACCCCTCCAGATATTTGGCCCTTT 1831
217 eRSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSer 233
1832 CTTCGCGAGTCTATTAATGGAACCTAACCTATCCCTTCCTTCGGAG 1881
234 AspProGlnArgLeuThrGlnLeuValGlnSerLeuMetPheSerHis 250
1882 GATCCCCAACCCCTCAGCGGGGTGTGTGATGATCCCTTATGTTCTTC 1931
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTh 267
1932 GCTACACTTGGGATGATTTGTCAACAGCTGCTGACAGACACTCTCACA 1981
267 lUGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
1982 AGGAGCGAGAGAGAAATCTGTTAGAGCTAGAAAAAATGTTCCCTGG 2031

```

284 AspGlyArgProThrGlnLeuGlnAsnGluLeuLeuAspMetGlyPheProle 300  
 2032 GACGGGCGACCCAGCTGCAATATGATGACATGAGATTTCCCTT 2081  
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGlySerLeuL 317  
 2082 GACTGCGCCCGGTGGACTACACACGCGTCAAGAGTGGAGAGACTTGA 2131  
 317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgAr 333  
 2132 AAATCTATGCGCAGGCTCTGGTGGCGGCTCTCCGGGGCGCTCAAGACGG 2181  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 2182 CCCACTAATTTGGCTAAGGTAAGAGAAATGATGACGAGGACCGAATGAAC 2231  
 350 GProSerValPheLeuGluArgLeuMetGlyAlaPheArgArgPheThr 367  
 2232 CCCCTCTGTTTCTTGGAGAGGCTCTTGGAGGCTTCAAGGGGTACACC 2281  
 367 rOpheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 2282 CTTTGTATCCACCTCAGAGGCGCCAAAGCCTCAGTGGCTTTGGCTTT 2331  
 384 lIleGlyInSerAlaLeuAspIleArgGlyLysLeuGlnArgLeuGlu 400  
 2332 ATAGGACACTCAGCCTTGATTTAGAAAGAAAGCTTCAGAGCTGGAAGG 2381  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417  
 2382 GTTACAGGAGGCTGAGTTACGTGATCTAGTGAAGGAGGACGAGAAAGAT 2431  
 417 yTfTyrArgArgGluThrGluGluGluLysGluGlnArgGlyLysGlu 433  
 2432 ATTACAAAGGAGACAGAAAGAAAGAAAGAAAGAAAGAAAGAGACAA 2481  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 2482 AGAGAGAAAGGAGGAAAGAGCTAATAACCGCAAGAGAAAGATTTGAC 2531  
 450 rLysIleLeuAlaIleValAlaGluGlyLysSerSerArgGluArgGlu 467  
 2532 TAAGATCTTGCTGCTGCTGTAAGGAGGAAAGCAATACGGAAGAGAGA 2581  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 2582 GAGATTTTGAAGAAATTTAGTCAGGCGCTCAGACAGTCAGGGAACCTGG 2631  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG 500  
 2632 AATAGGACCCCACTGACAGAGCAACATGTCATTTGTAAGAAAGG 2681  
 500 yHISrPalAArgAsnGlyProLysLysGlyAsnLysGlyProLysVal 517  
 2682 ACACCTGGCAAGAACTGCCCAAGAGGAAACAAAGAGCTGAAGGCT 2731  
 517 euAlaLeuGluGluAspLysAsp 524  
 2732 TAGCTCTGGAAGAGATTAAGAC 2754  
 seq\_name: gb\_pat:AX052647  
 seq\_documentation\_block: 6076 bp DNA PAT 12-JAN-2001  
 LOCUS AX052647  
 DEFINITION Sequence 32 from Patent WO001726.  
 ACCESSION AX052647  
 VERSION AX052647.1 GI:12226837  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus.  
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 REFERENCE 1 (bases 1 to 6076)

AUTHORS Federpiet, M.J.  
 TITLE Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 JOURNAL Patent: WO 00/1726-A 32 30-NOV-2000;  
 Mayo Medical Ventures (US)  
 FEATURES Location/Qualifiers  
 source 1..6076  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 1619 a 1505 c 1598 g 1354 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 2659.00 Length: 525  
 Ratio: 5.143 Gaps: 1  
 Percent Similarity: 98.476 Percent Identity: 95.619  
 alignment\_block:  
 US-09-171-553b-4 x AX052647  
 Align seg 1/1 to: AX052647 from: 1 to: 6076  
 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrp 17  
 1183 ATGGGACAGACGGTGACGACCCCTTGAATTGACTCTGACCATTTGGAC 1232  
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGly 34  
 1233 TGAAGTTAAACCCAGGGCTCATTAATTTGTCAGTTCAGTTAAGAGGAG 1282  
 34 rOTrPGLnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
 1283 CTTGGCAGACTTCTGTCTCTCAATGCGGACATTCGATTCGATTTGATGG 1332  
 51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAla 67  
 1333 CCATCAGAGGGGACCTTAATTCGAGATTAATCTGCTGTTAAAGCAAT 1382  
 67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
 1383 TATTTTTCAGCTGACCGCGGCTCTCATTCCTCAATCAGAGAGCCCTATTC 1432  
 84 eUThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100  
 1433 TTACGTGGCAGATTTGGCAGAGGATCTCCGCCATGAGTTAAACCTTGG 1482  
 101 LeuAsnLysProArgLysProGlyProArgGlnLeuAlaLeuGlyGlu 117  
 1483 CTGAATTAAGCCCAAGAAAGCCGAGTCCCGAATTTCTGCTCTTGGAGAGAA 1532  
 117 sASnLysHisSerAlaGluLysValGluProSerSerTyrLeuPro 134  
 1533 AAACCAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATATCTTACCC 1581  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 1582 GAGATTTGAGAGCGCGCGCTTGGCCGGAACCCCAATCTGTCCCCAC 1631  
 151 ProLeuSerSerThrGlyCysGlyGluGlyThr.SerAlaProProGly 167  
 1632 CCCATATCTGGACAGGCTGCTGCGAGGAGCCCTCTGCCCTCTCTGAG 1681  
 167 lAProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
 1682 CTCGCGCGGTGAGAGGACCTCTGACGAGGACTCGAGCGGAGGCGGCC 1731  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgTrpTrp 200  
 1732 ACCCGGAGCGGACAGAGAGATGCGCATTTACCGCTGCGCCACGTAAG 1781  
 200 yProProMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217  
 1781 yProProMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217

```

1782 CCCTCCACACACGGGGGCCAAATTCAGCCCTCCAGTATTGGCCCTTTT 1831
217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
1832 CTTCTGACAGATCTATATATGTGAAACATAACCATCCCTTCTCGGGAG 1881
234 AspProGlnArgLeuThrGlyLeuValGlySerLeuMetPheSerHisGly 250
1882 GATCCCAACAGCCCTCAGCGGGTGGTGGAGTCCCTATGTTCTCTCACCA 1931
250 nProThrTrpAspAspCysGlnGlnLeuGlnLeuGlnThrLeuPheThrTrp 267
1932 GCTACTTGGGATGATTTGTCACACAGCTGCTGCAGACACTTTCACACCG 1981
267 LngLArgGlnArgGlyLeuLeuGlnLysValArgLysAsnValProGlyAla 283
1982 AGGACGAGAGAGAATTTCTGTAGAGGCTAGAAAAAATGTTCTCTGGGGCC 2031
284 AspGlyArgProThrGlnLeuGlnAsnGlyLeuAspMetGlyPheProLe 300
2032 GACGGGCGCCAGCCAGCTGCAAAATGATGACATGAGGATTTCCCTT 2081
300 uThrArgProGlyTrpAspTyrAsnThrAlaGlnLysArgLysSerLeu 317
2082 GACTGCCCCCGTGGGACTTCACACGCGCTGAAGGTAGGGAGACTTGA 2131
317 ySileTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
2132 AATCTATTCGCGAGGCTCTGGTGGGGGCTCTCCGGGGGCGCTCAAGAGG 2181
334 ProThrAsnLeuAlaLysValArgGlnValMetGlnGlyProAsnGlyPr 350
2182 CCCACTAATTTGGCTAAGTAAGAACTGATGCAGGAGCCGAATGAACC 2231
350 oProSerValPheLeuGlnArgLeuMetGlnAlaPheArgArgPheThrP 367
2232 CCCCTGTTTCTTGAAGAGCTCTTGAAGCCCTTCAGGCGGTACACC 2281
367 rOPheAspProThrSerGlnLysAlaGlnLysAlaSerValAlaLeuAlaPhe 383
2282 CTTTGGATCCACCTCAGAGGCCCAAAAGCTCAGTGGCTTGGCTTT 2331
384 lIleGlyLysSerAlaLeuAspLysLysLysLeuGlnArgLeuGlnGly 400
2332 ATAGACAGTCAGCTTGATATTAGAAAGAGCTTCAGACACTGAGAGG 2381
400 yLeuGlnGlnLysLysLeuArgAspLeuValArgGlnLysValT 417
2382 GTTACAGAGGCTGAGTACGTTAGTCTAGTGAAGAGGAGAGAAAGTAT 2431
417 yTyrArgArgGlyThrGlnGlnGlnLysGlnArgLysGlnLysGln 433
2432 ATTACAAAAGGAGACAGAAAGAAAGGAAAGCAAAAGAAAGAGAGAA 2481
434 ArgGlnLysLysGlnLysArgAspArgArgGlnLysLysLysLysLys 450
2482 AGAGAGGAAAGGAGGAAAGAGCTAATAAACGCAAGGAAAGAAATTTGAC 2531
450 rLysIleLeuAlaValAlaGlnLysSerSerArgGlnArgGln 467
2532 TAAATCTTGGCTGACGTGGTTGAAGGAAAGCAATACGGAAGAGAGA 2581
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlnLysLys 483
2582 GAGATTATTAGGAAATAGTCAAGCCCTAGACAGTCAAGGGAACCTGGGC 2631
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlnLysGly 500
2632 AATGAGACCCCACTCGACAAAGCAATGTGCAATATTGTAAGAAAGAG 2681
500 yHisTrpAlaArgAsnGlyProLysGlnLysGlnLysGlnLysGlnLys 517
2682 ACACTGGGCAAGAGACTGCCCCAAGAAAGGAAACAAAGGACTGAAAGTCT 2731

```

```

517 euAlaLeuGlnLysAspLysAsp 524
2732 TAGCTCTGGAAGAGATTAAGAC 2754

```

seq\_name: gb\_pat:AX052638

seq\_documentation\_block: LOCUS AX052638 7873 bp DNA PAT 12-JAN-2001

DEFINITION Sequence 23 from Patent WO0071726.

ACCESSION AX052638

VERSION AX052638.1 GI:12226828

KEYWORDS porcine endogenous retrovirus.

ORGANISM Porcine endogenous retrovirus.

REFERENCE 1 (bases 1 to 7873) Viruses: Retrovirus: Retroviridae: Mammalian type C

AUTHORS Federpspiel, M.J. Methods to inhibit infectious agent transmission during

JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;

FEATURES Location/Qualifiers

BASE COUNT 2200 a 1914 c 1961 g 1798 t

alignment\_scores: Quality: 2659.00 Length: 525

Percent Similarity: 98.476 Percent Identity: 95.619

alignment\_block: US-09-171-553B-4 x AX052638

Align seg 1/1 to: AX052638 from: 1 to: 7873

```

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrp 17
62 ATGGAGACAGCGGTGACAGACCCCTTGTAGTTGACTGCGACCATTTGAC 111
17 rGlnValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyPr 34
112 TGAAGTTAAATCCAGGCTCATTAATTTGTACGTTAGGAAAGGAGAC 161
34 rOTrpgLInThrPheCysAlaSerGlnTrpProThrPheAspValGlyTrp 50
162 CTGGGAGACTTCTGTCTGTCTGATGGCCGACATTCATGTTGATGATG 211
51 ProSerGlnGlyThrPheAsnSerGluIleLeuAlaValLysAla 67
212 CCATCAGAGGGGACCTTAACTTGAGATTATCTGCTGTTAAAGCAAT 261
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGlyProTyrIle 84
262 TATTTTCAAGCTGAGCCGCGCTCATCCCAATCAGGAGCCCTATATCC 311
84 eUTHrTrpGlnAspLeuAlaGlnAspProProProThrValLysProTrp 100
312 TTACGCGGCAAGATTGGCAGAGATCCCTCCGCAATGAGGTTAAACCTTG 361
101 LeuAsnLysProArgLysProGlyProArgLysLeuAlaLeuGlyGly 117
362 CTGAATTAAGCCAAAGAACGACGATCCCGAATTTGCGCTTGGAGAGAA 411
117 sAsnLysHisSerAlaGlnLysValGlnProSerSerSerTyrLeuPro 134
412 AAACAAACACTCGGCTGAAGAAAGTCAGGCCCTTC.TCCTCATATCTACCC 460

```

```

134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProthr 150
||||| ||||||| ||||||| ||||||| |||||||
461 GAGATTGGAGACCGCGGCTTGGCCGAGACCCCAATCTGTTCCGCCACC 510
||||| ||||||| ||||||| ||||||| |||||||
151 ProLeuSerThrGlyCysCysGlyGlyThrSerAlaProProGlyAla 167
||||| ||||||| ||||||| ||||||| |||||||
511 CCTTATCTGGACACAGGTGCTGCGAGGAGACCCCTCTGCTCTCTGAG 560
||||| ||||||| ||||||| ||||||| |||||||
167 lAProValAlaGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
||||| ||||||| ||||||| ||||||| |||||||
561 CTCCGGCGGTGAGGAGACCTGCTCGAGGAGCTCGAGCGGAGGCGCC 610
||||| ||||||| ||||||| ||||||| |||||||
184 ThrProGluArgThrAspGluIleAlaIleuProLeuArgThrTyrGly 200
||||| ||||||| ||||||| ||||||| |||||||
611 ACCCGGAGCGGACAGAGATCGACATTACCGCTGCGAGCTAGCG 660
||||| ||||||| ||||||| ||||||| |||||||
200 yProPromeTProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhes 217
||||| ||||||| ||||||| ||||||| |||||||
661 CCTCCCAACCGGGGGGCGCATTCAGGCCCTCCAGATTTGGCCCTTTT 710
||||| ||||||| ||||||| ||||||| |||||||
217 eTSerAlaAspLeuTyrAsnTyrPlyThrAsnHisProProPhesGlu 233
||||| ||||||| ||||||| ||||||| |||||||
711 CTCTCGAGATCTATATATGGAACCTAACCATCCCTTTCTCGAG 760
||||| ||||||| ||||||| ||||||| |||||||
234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisGly 250
||||| ||||||| ||||||| ||||||| |||||||
761 GATCCCAACGCCCTCACGGGGTGGTGGAGTCCCTTATCTTCTCACC 810
||||| ||||||| ||||||| ||||||| |||||||
250 nProThrTrpAspPaspCysGlnGlnLeuGlnThrLeuPheThrThrG 267
||||| ||||||| ||||||| ||||||| |||||||
811 GCTACTTGGAGATGATTGTCAACAGCTGCTGACACACTCTTCACACCG 860
||||| ||||||| ||||||| ||||||| |||||||
267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
||||| ||||||| ||||||| ||||||| |||||||
861 AGAGCGAAGAGAAATCTGTAGAGCTAGAAATAATCTCTGGGGCC 910
||||| ||||||| ||||||| ||||||| |||||||
284 AspGlyArgProThrGlnLeuGlnAsnGlnIleAspMetGlyPheProLe 300
||||| ||||||| ||||||| ||||||| |||||||
911 GACGGCGACGCCGACGAGTTGCAAAATGAGATTGACATGGGATTTCCCTT 960
||||| ||||||| ||||||| ||||||| |||||||
300 uThrArgProGlyTyrPaspTyrAsnThrAlaGluArgGluSerLeuL 317
||||| ||||||| ||||||| ||||||| |||||||
961 GACTCCGCCCGGCTTGGAGCTACAAACAGGCTGAAAGCTAGGAGAGCTTGA 1010
||||| ||||||| ||||||| ||||||| |||||||
317 yTleIleArgGlnAlaLeuValAlaGlyLeuArgGlyLysSerArgArg 333
||||| ||||||| ||||||| ||||||| |||||||
1011 AAATCTATCGCCAGGCTGTGGTGGGGTCTCCGGGGCGCTTCAAGACGG 1060
||||| ||||||| ||||||| ||||||| |||||||
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
||||| ||||||| ||||||| ||||||| |||||||
1061 CCCACTAATTTGGCTAAGTAAAGAGATGATGACAGGACCGAATGACCC 1110
||||| ||||||| ||||||| ||||||| |||||||
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgAlaPheThrP 367
||||| ||||||| ||||||| ||||||| |||||||
1111 CCCCTCTGTTTCTTGTAGAGGCTTGTGAAGCTTCAAGCGGTACACCC 1160
||||| ||||||| ||||||| ||||||| |||||||
367 rPheAspProThrSerGluAlaGlnLysLaserValAlaLeuAlaPhe 383
||||| ||||||| ||||||| ||||||| |||||||
1161 CTTTGTGTTCCACCTCAGAGGCCCAAAAGCTTCACTGGCTTTGGCCTTT 1210
||||| ||||||| ||||||| ||||||| |||||||
384 lIleGlyIleSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
||||| ||||||| ||||||| ||||||| |||||||
1211 ATAGGACAGTCAGGCTTGATATTAGAAAGAGCTTCAGAGACTGGAAG 1260
||||| ||||||| ||||||| ||||||| |||||||
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluValT 417
||||| ||||||| ||||||| ||||||| |||||||
1261 GTTACAGAGGCTGAGTTACGTATAGTAAAGAGGCGAGAGAAAGTAT 1310
||||| ||||||| ||||||| ||||||| |||||||
417 yTyrArgArgGluThrGluGluGluLysGluGlnArgLysGluLysGlu 433
||||| ||||||| ||||||| ||||||| |||||||
1311 ATTCAAAAGGAGACAGAAAGAAAGGAGACAAAGAAAGAGAGAGAA 1360
||||| ||||||| ||||||| ||||||| |||||||
434 ArgGluGluArgGluArgArgArgArgArgGlnGlnLysAsnLeuTh 450
||||| ||||||| ||||||| ||||||| |||||||

```

---

```

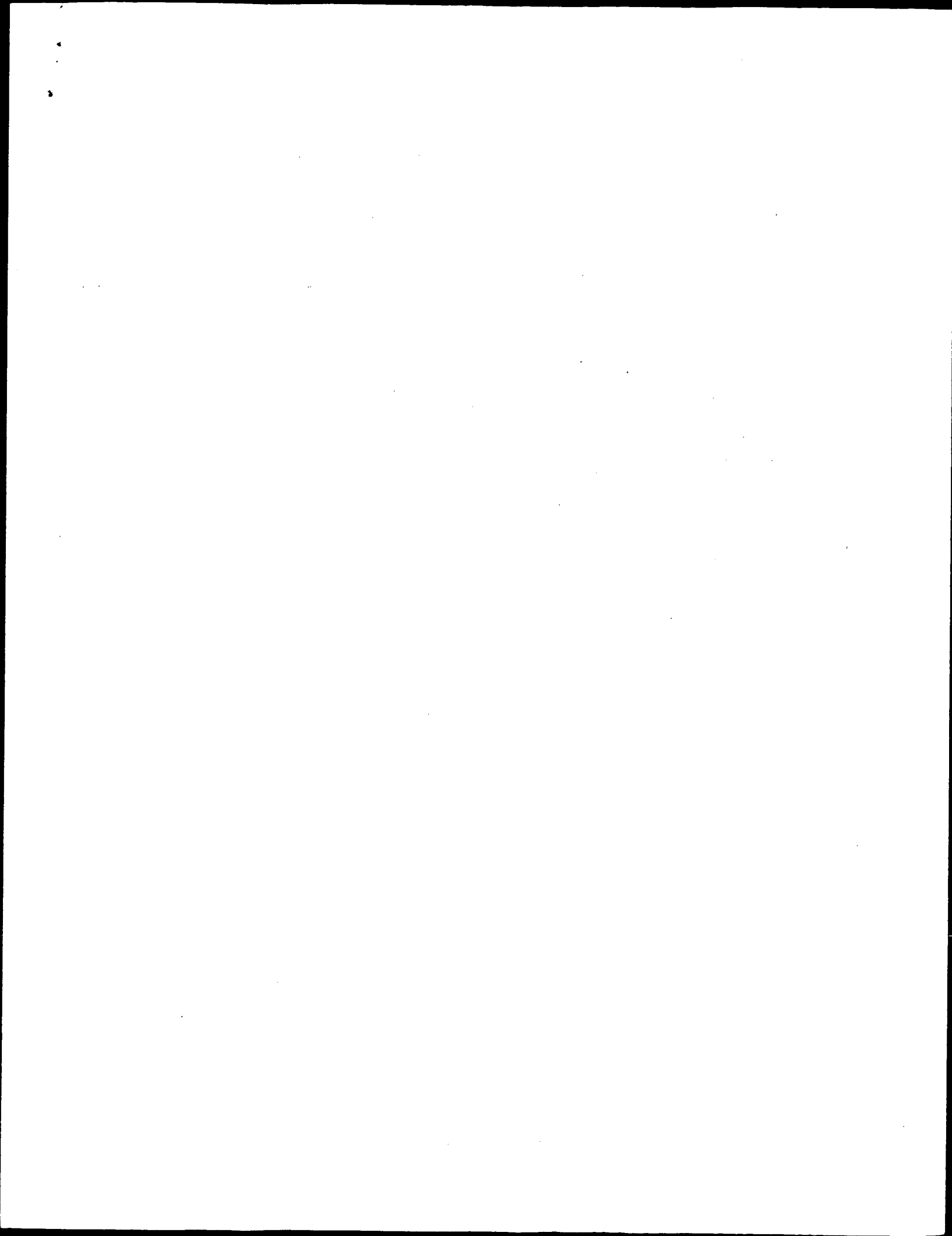
1361 AGAGAGGAAAGGAGGAGAAAGACGTAATTAACGGCAAGAAATTTGAC 1410
||||| ||||||| ||||||| ||||||| |||||||
450 rLysIleLeuAlaAlaValAlaGlyLysSerSerArgGluArgGluA 467
||||| ||||||| ||||||| ||||||| |||||||
1411 TAAATCTTGGCTGCACTGTGTTGAAGGAAACCAATTAACGAAAGAG 1460
||||| ||||||| ||||||| ||||||| |||||||
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
||||| ||||||| ||||||| ||||||| |||||||
1461 GAGATTTTAGGAAATTAATGATGAGCCCTTACAGATGACAGTACGGAAC 1510
||||| ||||||| ||||||| ||||||| |||||||
484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluVal 500
||||| ||||||| ||||||| ||||||| |||||||
1511 AATAGACACCCACTCGACAAAGACCAATGTGCAATATTGTAAGAAAAAG 1560
||||| ||||||| ||||||| ||||||| |||||||
500 yHisTTrpAlaArgAsnProLysGlyAsnGlyGlyProLysValL 517
||||| ||||||| ||||||| ||||||| |||||||
1561 ACCTGGGCAAGAGACTGCCCCAAGAGGAAACAAAGACTGTAGAGTCT 1610
||||| ||||||| ||||||| ||||||| |||||||
517 euAlaLeuGluGluAspLysAsp 524
||||| ||||||| ||||||| ||||||| |||||||
1611 TAGCTTGGAAGAAATTAAGAC 1633
||||| ||||||| ||||||| ||||||| |||||||
seq_name: gb_pat:AX052634
seq_documentation_block:
LOCUS AX052634 7362 bp DNA PAT 12-JAN-2001
DEFINITION Sequence 19 from Patent WO0071726.
ACCESSION AX052634
VERSION AX052634.1 GI:12226824
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
viruses; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1 Mammalian type C virus group.
1 (bases 1 to 7362)
REFERENCE
AUTHORS Federspiel,M.J.
TITLE Methods to inhibit infectious agent transmission during
xenotransplantation
JOURNAL Patent: WO 0071726-A 19 30-NOV-2000;
Mayo Medical Ventures (US)
FEATURES
Location/Qualifiers
source 1..7362
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1997 a 1821 c 1881 g 1663 t
ORIGIN
alignment_scores:
Quality: 2655.00 Length: 525
Ratio: 5.135 Gaps: 1
Percent Similarity: 98.476 Percent Identity: 95.429
alignment_block:
US-09-171-553b-4 x AX052634 ..
Align seg 1/1 to: AX052634 from: 1 to: 7362
1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTTrpTh 17
||||| ||||||| ||||||| ||||||| |||||||
1248 ATGGGACAGACGCTGACGACCCCTTGTAGTTGACTTCGACACATTGGAC 1297
||||| ||||||| ||||||| ||||||| |||||||
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34
||||| ||||||| ||||||| ||||||| |||||||
1298 TGAAGTTAATCCAGGCGTCATTAATTTGTCACTTCAGGTTAAGAGGAGC 1347
||||| ||||||| ||||||| ||||||| |||||||
34 rGTrpGlnThrPheCysAlaSerGluTTrpProThrPheAspValGlyTTrp 50
||||| ||||||| ||||||| ||||||| |||||||
1348 CTTGGGAGACTTCTGTCTCTGTAATGGCCGACATTCGATGTGGATGG 1397
||||| ||||||| ||||||| ||||||| |||||||
51 ProSerGluGlyThrPheAsnSerGluIleIleuAlaValLysAlaAl 67
||||| ||||||| ||||||| ||||||| |||||||

```

1398 CCATCAGAGGACCTTTAATCTGAGATTATCTGGCTGTTAAAGCAAT 1447  
67 eilepheglnthrglyproglyserhisproaspnglnuproyrilel 84  
|||||  
1448 TATTTTTCAGACTGACCGGCTCTCATGCCAATCAGAGCGCTTATCC 1497  
84 euthttrglnaspheualguaspuproproctripvalylsprottrp 100  
|||||  
1498 TTACGTGGCAGATTGGCAGAGATCTCCGCCCATGGTTAAACCTTGG 1547  
101 leuanlnysproarglysproglyproargyleuualaleuuglylly 117  
|||||  
1548 CTGATTAAGCCAGAAAGCCAGGTCCCGAATCTGGCTTGGAGAGAA 1597  
117 sasulysHiserialglulysvalgluproserSerSerTyrleuproa 134  
1598 AAACAACACTCGCTGAAAAAGTCAACCCCTC.TCCCATATCTACCCC 1646  
134 rgaspargglyalalaaspheualaglythrprothrCysSerprothr 150  
|||||  
1647 GAGATTGAGGAGCCCGCGCTTGGCCGGAACCCCAATCTGTTCCCCACC 1696  
151 ProleuSerSerThrGlyCysCysglnglnlythr.Serialproproglya 167  
|||||  
1697 CCTTATCTGGCACAGGCTGCTCGAGAGGACCTGCGCCCTCTGAG 1746  
167 laprovalvalglulyproalaaglythrargserargarglyala 183  
|||||  
1747 CTCCGCGCGTGAGGAGACCTGCTCGAGGACTCGAGCGGAGGCGGCC 1796  
184 Thrprogluarargthraspgluileualaileuuproleuarthrtrygl 200  
|||||  
1797 ACCCGGAGCGACAGAGAGATCGCAGCATTTACCGCTGCGCAGGTGCG 1846  
200 yproprometproglylglnglnleuuglnproleuuglnlytrtrprophes 217  
1847 CCTCCACACACCGGGGGGCCAATTGACGCCCTCAGTATTGGCCCTTT 1896  
217 erSerialaspheutryasntrplysThrasnHisPropropheserGlu 233  
1897 CTTCGTGAGATCTCTATAATTGGAAAACTAACCATCCCTCTCTCGAG 1946  
234 AspProgluarargleuthrglyleuvalgluSerleuemetpheserHisgl 250  
|||||  
1947 GATCCCAACGCCCTCAGCGGGTGGTGGAGTCCCTTATGTTCTCTCACA 1996  
250 nProthrTrpaspaspCysglnglnleuuglnThrleupheThrThrg 267  
1997 GCCTACTGGGATGATGTCACAGCTGCTGACAGACACTCTTCACAAACG 2046  
267 lugluarargluarglleuuleuglnualargylsasnvalProglyala 283  
|||||  
2047 AGGAGCGAGAGAAATCTGTAGAGGCTAGAAAAAATGTTCTCGGGGCC 2096  
284 AspGlyargProthrnglnleuuglnasngluileaspmetgilypheProle 300  
2097 GAGCGGGGACCCACCGCAGTTGCCAAATGAGATTGACATGGGATTTCCCTT 2146  
300 uThrargProglyTrpaspTyrasnThrleuaglyllyarggluSerleu 317  
2147 GACTCGCGCCGTTGGGACTACAAACAGGCTGAAGGTAGGAGAGAGCTTGA 2196  
317 ysiletyrarggluinalaleuualaglyleuargglyalaserargarg 333  
|||||  
2197 AAATCTATCCCGACGCTGTGTGGCGGCTCTCGGGCGCCCTCAAGACGG 2246  
334 ProThrAsnleuualalysvalarggluvalMetglnlyProasnglupr 350  
|||||  
2247 CCCACTAATTTGGCTAAGTAAAGAGATGATCAGGACCGGATATGAAACC 2296  
350 opProSerValPheleuugluargleuMetgluialapheargargpheThr 367  
|||||  
2297 CCCCCTCTGTTTCTTGAGAGGCTTTGGAAGCCTTCAGGCGGTACACCC 2346  
367 ropheaspProthrSerGluuaglnulysAlaSerValalaleuAlaPhe 383  
2347 CTTTGTATCCCACTCAGAGGCCCAAAAGCCTCAGTGGCTTGGCTTTT 2396  
384 IleGlylnSerAlaleuAspIleargylslysleuuglnargleuugln 400  
2397 ATAGACAGTCAGCTTGATATTAGAAAGACCTTCAGGACTGGAAG 2446  
400 yleuuglnluaglualeuargaspheuvalarggluualaglyllyval 417  
2447 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAAGAGGACAGAGAAATAT 2496  
417 YTrpArgarggluThrGluGluGlnlysglnarglnargylsGlu 433  
2497 ATTACAAAAGGAGACAGAGAAAGAAAGGACAAAGAAAAAGAGAGAA 2546  
434 ArgGluGluArgGluuargargaspargargGlnGluysasnleuth 450  
2547 AGAGAGGAAAGGAGGAAAGACCTAATAAACGCAAGAGAGAAATTTGAC 2596  
450 rlyslleuualalaValalaglylysSerSerArggluargGlu 467  
2597 TAAGATCTTGGCTGCAGTGTTGAAAGGAAAGCAATACGAAAGAGAGA 2646  
467 rgaspPhearglyslleargserGlyProargGlnSerGlyasnleuGly 483  
2647 GAGATTTTGAAGAAATTAGTCAAGCCCTTAGACAGTCAAGGAAACCTGGGC 2696  
484 AsnArgThrProleuaspLysaspGlnCysalaTyrCyslysGluysgl 500  
2697 AATAGAGACCCCACTCGACAAAGGACCAATGTGCATATTGTAAAGAAAAAG 2746  
500 yHisTrpAlaargAsnCysprolyslslyslslnlysglyProlyval 517  
2747 ACAGTGGCAAGGAACTGCCCAAGAGGAAAGGAAAGAAAGAGTGAAGTCT 2796  
517 euAlaleuGluGluaspLysasp 524  
2797 TAGCTGAAAGAGATTAAGAC 2819







OM of: US-09-171-553B-5 to: GenEmbl:\* out\_format : pfs  
 Date: Feb 24, 2002 10:10 AM  
 About: Results were produced by the GenCore software, version 4.5.  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODE=frame+p2n.model -DEV=xlh  
 -O:/cgn2.1/uspro\_spool/US09171553/tunat\_22022002.122825.11149/app\_query.fasta\_1.3312  
 -DB=GenEmbl -OFMT=fastap -SUFFIX=feb22std.rge -CAPOP=12.000  
 -CAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62  
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NORM-ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09171553 -CGN1\_1\_13817 -NCPU=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-171-553B-5  
 Query length: 1194  
 Database: GenEmbl:\*  
 Database sequences: 1472140  
 Database length: 341344837  
 Search time (sec): 5853.310000

## score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:A66553	+ 6304.00	6651.55	0.0	8209	! A66553 Sequence 3 from Patent
gb_pat:A66552	+ 6284.00	6630.40	0.0	8196	! A66552 Sequence 2 from Patent
gb_vi:PEN133816	+ 6232.50	6575.11	0.0	8918	! AJ133816 Porcine endogenous re
gb_vi:PEN293657	+ 6229.50	6572.09	0.0	8763	! AJ133657 Porcine endogenous re
gb_vi:PEN133817	+ 6223.50	6567.77	0.0	8849	! AJ133817 Porcine endogenous re
gb_vi:PEN133818	+ 6223.50	6565.76	0.0	8750	! AJ133818 Porcine endogenous re
gb_vi:PERV17013	+ 6202.50	6544.59	0.0	7808	! Y17013 Porcine endogenous reti
gb_vi:PEN293656	+ 6142.00	6479.33	0.0	8918	! AJ293656 Porcine endogenous ty
gb_pat:AR130475	+ 6099.50	6435.21	0.0	8132	! AR130475 Sequence 3 from patent
gb_pat:AX052635	+ 6095.50	6436.69	0.0	4402	! AX052635 Sequence 2 from pate
gb_pat:AX052634	+ 6095.50	6431.90	0.0	7362	! AX052634 Sequence 19 from Pate
gb_om:AF038600	+ 6087.50	6422.51	0.0	8132	! AF038600 Sus scrofa porcine en
gb_om:AF038599	+ 5979.50	6308.21	0.0	8132	! AF038599 Sus scrofa porcine en
gb_pat:AX052638	+ 5862.00	6184.16	0.0	7873	! AX052638 Sequence 23 from Pate
gb_pat:AX052647	+ 5846.50	5958.51	0.0	6076	! AX052647 Sequence 32 from Pate
gb_pat:AR130473	+ 5637.50	5946.35	0.0	8060	! AR130473 Sequence 1 from Pate
gb_pat:AX052636	+ 5635.50	5946.87	0.0	6076	! AX052636 Sequence 21 from Pate
gb_pat:A66551	+ 4840.00	5110.60	2.2e-276	3320	! A66551 Sequence 1 from Patent
gb_om:AF038601	+ 4899.50	4732.29	2.5e-255	7333	! AF038601 Sus scrofa porcine en
gb_pat:AR130474	+ 4489.50	4732.29	2.5e-255	7333	! AR130474 Sequence 2 from patent
gb_htg:AC092712	+ 4433.50	4641.60	2.9e-250	214764	! AC092712 Mus musculus clone
gb_htg:AC069310	+ 4424.50	4633.83	7.8e-250	177788	! AC069310 Mus musculus chromc
gb_htg:AC068810	+ 4424.50	4633.41	8.2e-250	185872	! AC068810 Mus musculus chromc
gb_htg:AL591952	+ 4418.50	4626.03	2.1e-249	207677	! AL591952 Mus musculus chromc
gb_htg:AC079547	+ 4412.50	4620.58	4.2e-249	188530	! AC079547 Mus musculus clone
gb_htg:AC079473	+ 4412.50	4619.89	4.6e-249	203155	! AC079473 Mus musculus clone
gb_ro:AC005743	+ 4378.00	4584.22	4.5e-247	185548	! AC005743 Mus musculus chromc
gb_pat:AC093350	+ 4374.00	4579.44	8.3e-247	196755	! AC093350 Mus musculus chromc
gb_om:AF151794	+ 4359.00	4592.88	1.5e-247	8431	! AF151794 Phascollartectos clne
gb_htg:AC080144	+ 4357.00	4561.19	8.6e-246	202333	! AC080144 Mus musculus chromc
gb_htg:AL592224	+ 4355.00	4558.34	1.2e-245	218938	! AL592224 Mus musculus chromc
gb_ro:MHHC438N12	+ 4344.50	4550.75	3.3e-245	149886	! AF049850 Mus musculus major
gb_pat:AR136169	+ 4339.50	4572.00	2.2e-246	8655	! AR136169 Sequence 1 from patent
gb_vi:AF053745	+ 4339.50	4572.00	2.2e-246	8655	! AF053745 Mus dunni endogenous
gb_htg:AC073736	+ 4337.50	4540.67	1.2e-244	199883	! AC073736 Mus musculus clone
gb_ro:AC012147	+ 4331.00	4533.02	3.2e-244	217034	! AC012147 Mus musculus, clone
gb_vi:GLJ060065	+ 4325.00	4556.96	1.5e-245	8379	! U60065 Gibbon leukemia virus s
gb_htg:AC013797	+ 4320.00	4521.28	1.4e-243	219294	! AC013797 Mus musculus clone
gb_vi:PCGGPE	+ 4319.00	4550.94	3.2e-245	8088	! M26927 Gibbon ape leukemia vir
gb_htg:AC074337	+ 4317.00	4523.02	1.2e-243	129317	! AC074337 Mus musculus chromc

gb\_htg:AC073758  
 gb\_htg:AC091703  
 gb\_htg:AC079555  
 gb\_htg:AC087145  
 gb\_hcg:AC020959

seq\_name: gb\_pat:A66553

seq\_documentation\_block:

LOCUS A66553 8209 bp DNA PAT 29-MAR-1999  
 DEFINITION Sequence 3 from Patent WO9740167.  
 ACCESSION A66553  
 VERSION A66553.1 GI:4538106  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus

Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 8209)  
 Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.

PORCINE RETROVIRUS  
 Patent: WO 9740167-A 3 30-OCT-1997;  
 Q ONE BIOTECH LTD (GB)

FEATURES  
 source  
 1..8209  
 .Location/Qualifiers  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 588..2162  
 /codon\_start=1  
 /product="VIRAL CORE PROTEIN"  
 /protein\_id="CAB39347.1"  
 /db\_xref="GI:4538107"

translations="MSGVTVTPLSLTDHWTEVRSRAHNLISVQKKGQTCASEMP  
 TFDGPPSEGTNSEILAVKALIFQGTGSDPDQEPYILTWDLAEDPPWVKPWLN  
 KPRKPGRIALGKNNKSAEKVESSYLPRDGDAGLAGTPTCSPTLSPPVCCGEG  
 TSAPPAGVVEGPAATRRGATPERTDEAILPLRTYPPPGGOLAPLOYPFSS  
 ADLYNKNTHPPSEDPQRLTGIVESIMFESHQPTWDCQQLQTLFTTERERILLEA  
 RKNVPGADGPTOLQEDIMGFPLTRPGWYNTAEGRESLKIYRQALVAGLGRASRRP  
 TNLAKYEVNMGQNPSPVSLRMEAFRRPTPDTSQAQKASVALATIGQSALDIR  
 KKLQRLGLEQAEERDLVREAEKYVYRTEEEKERKEKERERERERDRQKELNLT  
 KILAAVVEGKSSRREREDFRKIRGSPQSGNLGNRLPLDKDQACAYCKEKGHWARNCPK  
 <2163..5747  
 /codon\_start=1  
 /product="REVERSE TRANSCRIPTASE"  
 /protein\_id="CAB39348.1"  
 /db\_xref="GI:4538108"

CDS

translations="GRRGSDPLPEPRVTLKVEGQVFEVLDTGAESHVLLQPLGLKLE  
 KSWMGATGQROYPMTRTVDLGVGRVTHSPFVPECPVLLGRDLTKGAQISF  
 EOGRPVSVNNKPIVTILQLDDEYRLSYPQKPDODIOSWLEOPQANETAGMGLA  
 KOVPPQVQLKASATPSVRYQPLSREARGHPVQRILIQOGLIVPVSPWNTPLLP  
 VRKPGNDYRPVODLREVNKRKODIHTPTPNPNLLSALPPERNWTVYLDLKDAPFLC  
 RLHTQSPQFAFWRDPGTGRTQLTWRLPQKNSPTIFDEALHRDLANFRIOHPQ  
 VTLQYVDDLLLAGATKQDCEGTKALLLSLDGLYRASAKAQICRRREVYLSLR  
 GQSWLTERAKFTVQIPATPAVALPQREFLGATGFCRLWPGFATLAAPLYPLTKEGK  
 GFSWAEHOKAFKAIKKALISAPALPDVTKPFTLYVDERGVARGVLTQLPWRR  
 PVAYLSKKLDPVAGWPVCLKAATAVAAILVKDADKLITLQNTIVTAPHALENIYQPP  
 DRWTHNMTYVOSLLTDERVTPFAALNPALPPEDEPVDHCDHLLLEETGRV  
 KLTLDPLTGEVITFTDSSVYVEGKRMAGAAVDGTTRTINASSPEGTSQAQALM  
 ALFQALRAEGKNTINTYRSYAFATAHVGAIKQGLLTAGREIKNKEITLSLE  
 ALHLKRLAIKHCQGHQKADLLSRGNOMADRVAKAQAVALPIETPKAPEPRQ  
 LITLQWQETIKIDQSETPEGTCTYSGYKEILPKHGELEYVQOIHRLTHLTKHLOQL  
 VRTSPYHVLRLPQADSVVHKVCQQLVNPANSPRIPGRKLRGSHPGAWEDVTFQK  
 PAKYGNKYLIVVDTSWGVNAPTRKSTVYVAKKILIEIFRFGIPGSDNGPQA  
 FVAQVSGGLAKILIGDKWLHCAYRQSSPOGVERMNETIKETIKTLTKTETGINWALL  
 PFVLRNTPGPGFTYELLYGPPPLAEAFSAHSDVLLSQPLSRKLEWVRQ  
 RANKQLREAYSGDLOVPHRFQVDSVYVRRHAGNLETWKGPYLVLTTPATKVK  
 GIPTLWASHVKKPAPPPDSQWRAEKENPLKRLHLRVPSNNNSPGQ"  
 5620..7590  
 /codon\_start=1  
 /product="VIRAL ENVELOPE PROTEIN"  
 /protein\_id="CAB39349.1"  
 /db\_xref="GI:4538109"

CDS

translations="GRRGSDPLPEPRVTLKVEGQVFEVLDTGAESHVLLQPLGLKLE  
 KSWMGATGQROYPMTRTVDLGVGRVTHSPFVPECPVLLGRDLTKGAQISF  
 EOGRPVSVNNKPIVTILQLDDEYRLSYPQKPDODIOSWLEOPQANETAGMGLA  
 KOVPPQVQLKASATPSVRYQPLSREARGHPVQRILIQOGLIVPVSPWNTPLLP  
 VRKPGNDYRPVODLREVNKRKODIHTPTPNPNLLSALPPERNWTVYLDLKDAPFLC  
 RLHTQSPQFAFWRDPGTGRTQLTWRLPQKNSPTIFDEALHRDLANFRIOHPQ  
 VTLQYVDDLLLAGATKQDCEGTKALLLSLDGLYRASAKAQICRRREVYLSLR  
 GQSWLTERAKFTVQIPATPAVALPQREFLGATGFCRLWPGFATLAAPLYPLTKEGK  
 GFSWAEHOKAFKAIKKALISAPALPDVTKPFTLYVDERGVARGVLTQLPWRR  
 PVAYLSKKLDPVAGWPVCLKAATAVAAILVKDADKLITLQNTIVTAPHALENIYQPP  
 DRWTHNMTYVOSLLTDERVTPFAALNPALPPEDEPVDHCDHLLLEETGRV  
 KLTLDPLTGEVITFTDSSVYVEGKRMAGAAVDGTTRTINASSPEGTSQAQALM  
 ALFQALRAEGKNTINTYRSYAFATAHVGAIKQGLLTAGREIKNKEITLSLE  
 ALHLKRLAIKHCQGHQKADLLSRGNOMADRVAKAQAVALPIETPKAPEPRQ  
 LITLQWQETIKIDQSETPEGTCTYSGYKEILPKHGELEYVQOIHRLTHLTKHLOQL  
 VRTSPYHVLRLPQADSVVHKVCQQLVNPANSPRIPGRKLRGSHPGAWEDVTFQK  
 PAKYGNKYLIVVDTSWGVNAPTRKSTVYVAKKILIEIFRFGIPGSDNGPQA  
 FVAQVSGGLAKILIGDKWLHCAYRQSSPOGVERMNETIKETIKTLTKTETGINWALL  
 PFVLRNTPGPGFTYELLYGPPPLAEAFSAHSDVLLSQPLSRKLEWVRQ  
 RANKQLREAYSGDLOVPHRFQVDSVYVRRHAGNLETWKGPYLVLTTPATKVK  
 GIPTLWASHVKKPAPPPDSQWRAEKENPLKRLHLRVPSNNNSPGQ"  
 5620..7590  
 /codon\_start=1  
 /product="VIRAL ENVELOPE PROTEIN"  
 /protein\_id="CAB39349.1"  
 /db\_xref="GI:4538109"

CDS

/translation="MHPTLSRRHLPTTRGGEKRRRLIPUSFASIANWFLTLTITPQASSK  
RLDSSNHRPLSLTLWLIIDPDGTGVNSTRGVAPRGVTWMPHELHFLRLINPAVKSTP  
PNLVRSYGYCCPGTEKERYKSGSGESFCRRWSCVTSNDGDKWPIISLDQDRVKFSEVN  
SGPGKYKMKMLYKDCSPSDDLKISFTRKTKGYSKVDKWYELGNSFLLYGGAG  
STLTIRLRIETCTEPPVAMPQDKVLAEOGPPALEPHNLVPOLTSIRPDITQPPSNS  
TTGLIINTPRNSGPVPTGQRLFSLIQOAFQAINSTDPTATSSCWLCILSSGPPYYE  
GMAKERFNTWEHRNOCWTGSRNKLITLTVSGKGTICGKAPPSSHOLICYSTVWYEQ4  
SENYILVPGYNRMWACNTGLTFCVSTSVENQKDFCVWQVIIPRVYIHPPEVYLDYD  
YRNRPKPESLTLAVMLGLGTAGVGTGAALITGPOOLEKGLGELHAAMTEDLRA  
LKESVNLEESLTSLSSEVLQNRGLDLFLREGGLCAALKECCCFYVDHSGAIRDMS  
NKLKRLERREREADOCWFSGWNRSPWMTLLSALTGLPLVLVLLLLLTITVGPCLINR  
FNAPYRSVAQIMVLRQOYGLLSQGETDL"  
BASE COUNT 2168 a 2064 c 2154 g 1823 t  
ORIGIN

alignment\_scores:  
Quality: 6304.00 Length: 1194  
Ratio: 5.280 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-171-553b-5 x A66553

Align seg 1/1 to: A66553 from: 1 to: 8209

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
2163 GGGAGACGGGGTTCGGACCCCTCCCGAGCCGAGGTAACCTTTGAAGGT 2212  
17 IGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHiserv 34  
2213 GGAGGGGAACACAGTTGAGTTCCTGGTTGATACCGGAGGAGCATTCAG 2262  
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
2263 TCCTCTCAACACATTAGAAACTTAAAGAAATAAAATCTCGGTGATG 2312  
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValas 67  
2313 GGTGCCACAGGCAACGGCAGTATCCATGGACTACCGGAAGAACCGTTGA 2362  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84  
2363 CTTGGGAGTGGGACGGGTAAACCACTCGCTTCTGGTCACTCCCTGAGTGC 2412  
84 roValProLeuLeuGlyArgAspLeuThrLysMetGlyAlaGlnIle 100  
2413 CAGTACCCCTTCTAGGTAGACATTACTACCAAGATGGGAGCTCAAAAT 2462  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
2463 TCTTTTGAACAGGAAGACCAAGAGTGTCTGTGAATAACAACCATCAC 2512  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
2513 TGTGTGACCTCCCAATTAGATGATGATATCGACTATATCTCCCCCAAG 2562  
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
2563 TAAACCTGTATCAAGATATACAGTCTCGTTGGAGCAGTTTCCCCCAAGCC 2612  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
2613 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAGGT 2662  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
2663 TATTCAACTGAAGGCCAGTGTGTACACCATGATCATGTCAGACAGTACCCCT 2712  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
2713 TGAGTAGAGAGGCTCGAAGGAATTTGGCCGATGTTTCAAGAGTTAATC 2762

201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
2763 CAACAGGGCATCTAGTCTCTGTCCAATCCCTTGAATACTCCCTGCT 2812  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
2813 ACCGGTTAGAACCCCTGGGACCAATGATTATGACACAGTACAGGACTGA 2862  
234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
2863 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCGGAACCT 2912  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrVal 267  
2913 TATAACCTCTTGAAGCCCTCCCGCTGAAACGGAACCTGGTACACAGTAT 2962  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
2963 GGACTTAAAGATGCTCTTCTGCTGAGATTACACCCCACTAGCCCAAC 3012  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
3013 CACTTTTTCGCTTCAATGGAGAGATCCAGGTACGGGAAGAACCCGGCAG 3062  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
3063 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3112  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3113 TGACGAAGCCCTACACAGGGACCTGGCCAACTTCAGAGTCCCAACACCTC 3162  
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
3163 AGGTGACCTCTCCAGTACGTGAGTACCTGCTCTGGCGGGAGCAC 3212  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
3213 AAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAATGTCTGA 3262  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3263 CCTAGGTACAGAGCCTCTGCTAAGAGGCCAGATTTGCAGGAGAGAGG 3312  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
3313 TAACATACTTGGGTACAGTTTTCGGGGGGGGCAGCATGGCTGACGGAG 3362  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417  
3363 GCACGGAAGAAAACTGTAGTCCAGATACCGGCCCAACCAACAGCCAAACA 3412  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
3413 AGTGAGAGAGTTTTGGGGACACCTGGATTTTGCACACTGTGGATCCCGG 3462  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
3463 GGTTTGGACCTTAGCAGCCCTCTACCCGCTAACCAAGAAAAAGGG 3512  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysly 467  
3513 GGATTCCTCTGGGCTCTGTAGACACCAAGAGGCATTTGATGCTATCAAAA 3562  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
3563 GGCCCTGCTGAGCGACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTT 3612  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
3613 TTACCTTTTATGTGGATGACGCTAAGGGAGTAGCCCGGAGGAGTTTAACC 3662

501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
3663 CAAACCTAGGACCTAGGAGAGACTGTGCCTACCTGTCAAGAAGCT 3712  
|||||  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
|||||  
3713 TGATCCTGTAGCCAGTGGTGGCCGTAIGTCTGAAGGCTATCGCAGCTG 3762  
|||||  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
3763 TGCCCATACTGCTCAAGGACGCTGACAAATTGACTTTGGGACAGAAATA 3812  
|||||  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
3813 ACTGTAATAGCCCCCATGCTATGGAGAACATCGTTGCGCAGCCCCCAGA 3862  
|||||  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
|||||  
3863 CCGATGGATGACCAACGCCCGCATGCCCACTATCAAAAGCCTGTCTCTCA 3912  
|||||  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
3913 CAGAGGGGTCACTTTCGCTCCACCGCGCTCTCAACCCCTGCCACTCTT 3962  
|||||  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
3963 CTGCCTGAAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 4012  
|||||  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4013 GATTGAGGAGACTGGGGTCCGACGAGGACCTTACAGACATACCGCTGACTG 4062  
|||||  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4063 GAGAAGTCTAACCTGGTTCTACTGACGGAAGCAAGTATGTGGTGAAGGT 4112  
|||||  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4113 AAGAGGATGGCTGGGGCGGACGTGTGACGGGACCCGACGATCTGGGC 4162  
|||||  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluIleuMetAlaL 684  
|||||  
4163 CAGCAGCCTGCGGAAGAACTTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4212  
|||||  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4213 TCACGCAAGCTTTCGGGCTGGCCGGAAGGAAATCCATAACATTTATACG 4262  
|||||  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4263 GACAGCAGTATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAA 4312  
|||||  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4313 ACAAAAGGGGTGGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4362  
|||||  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4363 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4412  
|||||  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
4413 ATACACTGTCTGGACATCAGAAAGCCCAAGATCTCATATCTAGAGGGA 4462  
|||||  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
|||||  
4463 CCAGATGGCTGACCGGGTGGCAAGCAGCGACGCCAGGCTGTATACCTTC 4512  
|||||  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
4513 TCCCTATATAGAAACGCCCAAGCCCAAGAACCCAGACCCAGACGATACACC 4562  
|||||  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||

4563 CTAGAAACACTGGCAAGAGATAAAAGATAGACCAGTTCTCTGAGACTCC 4612  
|||||  
817 oGluGlyThrCystTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
4613 GGAGGGACCTGCTATACCTCATATGGAAGAAATCTCTGCCCCACAAAG 4662  
|||||  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
4663 AAGGTTTAGAATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 4712  
|||||  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
4713 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 4762  
|||||  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
4763 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATTGTGTGCCCTGCCAGCTGG 4812  
|||||  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
4813 TTAATGCTAATCTTCCAGAAATACCTCCAGGAAAGAGACTAAGGGGAAGC 4862  
|||||  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
4863 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCGCGCTAA 4912  
|||||  
917 sTyrGlyValAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
4913 ATACGGAACAATATCTATTGGTTTGTGTAGACACCTTTTCAGATGGG 4962  
|||||  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
4963 TAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5012  
|||||  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
5013 ATACTGGAGGAAATTTTCCAAAGATTGGAAATACCTTAAGGTAATAGGGTC 5062  
|||||  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5063 AGACAAATGGTCCAGCTTTCGTGCCCCAGGTAAAGTCAGGACTGGCCNAGA 5112  
|||||  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
5113 TATTGGGSAITGATTGGAACACTGCATTTGCATACAGACCCCAAGCTCA 5162  
|||||  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
5163 GGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGAGACCTTACCATAAT 5212  
|||||  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
5213 GACCACAGAGACTGGCATTAAATGATGGATGGCTCTCCTGCCCTTTGTGC 5262  
|||||  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
5263 TTTTTAGGCTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5312  
|||||  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5313 TTGCTCTACGGGGGACCCCGCCCTTGGCAGAAATTTGCCATAG 5362  
|||||  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
5363 TGCTGATGTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCGCTCG 5412  
|||||  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
|||||  
5413 AGTGGGTGAGGCAGCGGTGGNAGCAGCTCCGGGAGGCTACTCAGGA 5462  
|||||  
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117  
|||||

5463 GGAGACTTGCAGTTCCACATCGCTTCCAAAGTTGGAGATTCAAGTCTATGT 5512  
 1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134  
 5513 TAGAGCCACCGTGCAGGAACCTCGAGACTCGGTGGAGGACCTTATC 5562  
 1134 euValLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150  
 5563 TCGTACTTTTGGACCAACACAGCGGTGTGAAAGTCAAGAAATCCCCACC 5612  
 1151 TrpIleHisAlaSerHisValLysProAlaProProProAspSerGlyTr 1167  
 5613 TGAATCATCATCCACGCTTAAGCGCGCCACCTCCGACTCGGGTG 5662  
 1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeu 1184  
 5663 GAGAGCGGAAAGACTGAGAATCCCTTAAGCTTCGCTCCATCGCTGG 5712  
 1184 aProTyrSerAsnAsnSerProGlyGln 1194  
 5713 TTCCTTACTCTAACAACTACTCCCGAGGCCAG 5744

seq\_name: gb\_pat:A66552

seq\_documentation\_block: 8196 bp DNA PAT 29-MAR-1999  
 LOCUS A66552  
 DEFINITION Sequence 2 from Patent WO9740167.  
 ACCESSION A66552  
 VERSION A66552.1 GI:4538105

KEYWORDS porcine endogenous retrovirus.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM  
 Viruses; Retroviridae; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.

REFERENCE  
 1 (bases 1 to 8196)  
 Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

AUTHORS  
 TITLE PORCINE RETROVIRUS  
 JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;  
 Q ONE BIOTECH LTD (GB)

FEATURES  
 Location/Qualifiers  
 source  
 1..8196  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"

BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others  
 ORIGIN

alignment\_scores:  
 Quality: 6284.00 Length: 1194  
 Ratio: 5.272 Gaps: 0  
 Percent Similarity: 99.832 Percent Identity: 99.749

alignment\_block:

US-09-171-553B-5 x A66552 ..

Align seg 1/1 to: A66552 from: 1 to: 8196

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
 2149 GGGGAGACGGGTTCGAGCCCTCCCGAGCCCGAGGTAACCTTTGAAGGT 2198  
 17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
 2199 GGAGGGCAACCAAGTTGAGTTCCTGTTGATACCGGAGCGGAGCATTCAG 2248  
 34 aLeuLeuGlnProLeuLysLeuLysGluLysLysSerTrpValMet 50  
 2249 TCGTCTACACCACTTAGGAAACTTAAAGAAAAAATCTCGGTGATG 2298  
 51 GlyAlaThrGlyGlnArgGlnTyrProThrArgArgThrValas 67  
 2299 GTGCGCACAGGCAACGCGAGTATCCATGGACTACCCCGAAGAACCGTTGA 2348

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
 2349 CTTGGGAGTGGGAGCGGTAAACCCACTCGTTTGGTTCATCCCTGAGTGCC 2398  
 84 roValProLeuLeuGlyArgAspLeuThrLysMetGlyAlaGlnIle 100  
 2399 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2448  
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
 2449 TCCTTTGAACAAGGACAGCCAGAGTGTCTGTGAATAACAACCCATCAC 2498  
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
 2499 TGTGTTGACCTCCCAATTAGATGATGAATATCGACTATATTTCTCCCAAG 2548  
 134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
 2549 TAAAGCCTGATCAAGATATACAGTCTCGTGGAGCAGTTTCCCAAGCC 2598  
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
 2599 TGGCAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 2648  
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
 2649 TATTCAACTGAAGCCAGTGTCTACACCACTATCAGTCAGACAGTACCCCT 2698  
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
 2699 TGAGTAGAGGCTCGAGAAAGGAATTTGGCCGCTGTTCAAGATTAAATC 2748  
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
 2749 CAACAGGGCATCTAGTTCCTGTCCTCAATCCCTTGGAAATCTCCCTGCT 2798  
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
 2799 ACCGGTTAGGAAGCCTGGGACCAATGATTATCGACCACTAGTACAGACTGA 2848  
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
 2849 GAGAGGTCAATAAAGGGTGCAGGACATACCCCAACGGTCCCGAACCCCT 2898  
 251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267  
 2899 TATAACCTCTTGAGCGCCCTCCCGCTGACGGAAGTGGTACAGATT 2948  
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
 2949 GGACTTAAAGATGCCCTTCTTCTGCTGAGATTACACCCCACTAGGCAAC 2998  
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
 2999 CACTTTTGGCTTCGAATGGAGATCCAGGTACGGGAAGAACCGGGCAG 3048  
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
 3049 CTCACCTGGACCGAGTCCCAAGGTTCAAGAACTCCCGGACCATCTT 3098  
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
 3099 TGACGAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAACACCCCTC 3148  
 334 lValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
 3149 AGGTGACCCCTCCAGTACGTAGTACCTGCTTCTGGCGGGAGCCACC 3198  
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
 3199 AACAGAGCTGCTTAGAAGGTACGAAGGCCTACTGCTGGAATTGCTGA 3248  
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluV 384

|||||  
3249 CCTAGGCTACAGCCCTCTGCTAAGAAAGGCCAGATTTCAGGAGAGG 3298  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTyrLeuThrGlu 400  
|||||  
3299 TACATACTTGGGTACAGTTTCGGGGCGGCGGACGATGCTGACGGAG 3348  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417  
|||||  
3349 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3398  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
3399 AGTCAGAGAGTTTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCGG 3448  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
3449 GGTITGGACCTTACGACGCCACTCTACCCGCTAACCAAGAAAGGG 3498  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
3499 GGATTCCTCTGGGCTCTGAGCACCAGAGGCATTTCATCAAAAA 3548  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484  
|||||  
3549 GGCCCTGCTGAGCGACCTGCTCTGGCCCTCCTGACGTAACATAACCCCT 3598  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
3599 TTACCTTTTNGTGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 3648  
501 GlnThrLeuGlyProTrpArgArgProValAlaThrLeuSerLysLysLe 517  
|||||  
3649 CAAACCCCTAGSACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAGCT 3698  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534  
|||||  
3699 TGATCCTGTAGCCAGTGGTGGCCGTATGCTGAAGGCTATCGCAGCTG 3748  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
3749 TGGCCATACATGGTCAAGAGCCCTGACAAATGACTTTGGGACAGATATA 3798  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
3799 ACTGTAATAGCCCCCATGCTATGGACAACATCGTTCCGCGAGCCCCAGA 3848  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
3849 CCGATGGATGACCAACCGCCGATGACCCACTATCAAAGGCTGCTCTCA 3898  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
3899 CAGAGAGGTCACTTTTCGCTCCACCAGCCCTCTCAACCCCTGCCACCTT 3948  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
3949 CTGCCTGAAGAGACTGATGAACAGTGAACATCATGATTGCCATCAACTAT 3998  
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
3999 GATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCCGCTGACTG 4048  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4049 GAGAACTGCTAACTGGTTCTGCTGACCGAAGACAGCTATGTGGTGAAGGT 4098  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4099 AAGAGGATGCTGGGGCGGAGTGGTGGACGGGACCCGCGCATCTGGGC 4148  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||

4149 CAGACACCTTCCCGAAGAGAACTTCAGCGCAAAAGGCTGAGCTCATGGGCC 4198  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4199 TCACGGAAGCTTTGGGCTGGCGAAGGAAATCCATAAACATTTATAGC 4248  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4249 GACAGCAGGTATGCCCTTTCGACTGCACACGTCACAGGGGCCATCTATAA 4298  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4299 ACAAGGGGTTGCTTTACCTCAGCAGGGAGGAAATAAAGAACAAAGAGG 4348  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4349 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4398  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
4399 ATACACTGCTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4448  
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784  
|||||  
4449 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGCTGTTAACCTTC 4498  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
4499 TGCCCTATATAGAAAGCCCAAGCCCAAGACCCAGACAGCAGTACACC 4548  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
4549 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGAGTCC 4598  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
4599 GGAGGGGACCTGCTATACCTCATATGGGAAGGAAATCTGCCCCACAAG 4648  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
4649 AAGGTTAGAAATATGTCACAGATACATCGTCTAACCCACCTAGGAAC 4698  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
4699 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTCTGAGGCT 4748  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
4749 ACCAGAGTGGCTGACTCGGTGTCAAACATTTGTGCGCCTGCCAGCTGG 4798  
884 aIAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
4799 TTAATGCTAATCCTTCCAGAATACCTCCAGGAAGAGACTAAGGGAAGC 4848  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
4849 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAA 4898  
917 sTyrGlyAsnLysTyrLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
4899 ATACGGAACAAATATCTATGTTTTTGTAGACACCTTTTCAGGATGGG 4948  
934 aIgluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
4949 TAGAGCCTTATCCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAGAAA 4998  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
4999 ATACTGGGGAATTTTTCCAAGATTTGGAATACCTAAGGTAATAGGGTC 5048  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5049 AGACAATGTCTCAGCTTTCTGTTGCCAGGTAAGTCAGGAGCTGGCCAAGA 5098



```

984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
|||||
5099 TATTGGGATTGATTGGAACCTGCTTGTGATACAGACCCCAAGCTCA 5148
|||||
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
5149 GGACAGGTAGAGAGATGAATAGAACCAATTAAGAGAGACCCCTACCAATT 5198
|||||
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||
5199 GACCACAGAGACTGCATTAATGATGGCTGCTCTCTGCCCTTTGTGC 5248
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5249 TTTTGGGTGAGGAACACCCCTGGACAGTTTGGCTGCACCCCTATGAA 5298
|||||
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
5299 TTGCTCTACGGGGGACCCCGCTGGCAGAAATTCCTTTGCACATAG 5348
|||||
1067 tAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
5349 TGCTGATGCTGCTTTCCAGCCTTTGTTCTAGGCTCAAGCGCTCG 5398
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
|||||
5399 AGTGGGTGAGGACGAGCGTGAAGACAGCTCCGGGAGGCTACTCAGGA 5448
|||||
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
|||||
5449 GGAGACTTGCAAGTTCCACATTCGCTTCCAGTTGGAGATTACAGTCTATGT 5498
|||||
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
|||||
5499 TAGAGCCACCGTCGAGAACTCGAGACTCGGTGGAGGACCTTATC 5548
|||||
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
|||||
5549 TCGTACTTTTACCACACCAACCGCTGTGAAGTGAAGGAATCCCCACC 5598
|||||
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
|||||
5599 TGGATCCATGATCCCGTAAAGCGCGCCACCTCCCGACTCGGGGTG 5648
|||||
1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
|||||
5649 GAGAGCCGAAAAGACTGAGATCCCTTAAGCTTCGCTCCATCGCCTGG 5698
|||||
1184 alProTyrSerAsnAsnSerProGlyGln 1194
|||||
5699 TTCCTTACTCTAACAATAACTCCCGAGGCCAG 5730
|||||

```

seq\_name: gb\_vi:PE133816

seq\_documentation\_block:

LOCUS PE133816 8918 bp DNA VRL 07-DEC-2000  
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env  
genes and LTR (class B, clone 33).

ACCESSION AJ133816

VERSION AJ133816.1 GI:6688945

KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol  
protein.

SOURCE

ORGANISM

porcine endogenous retrovirus.  
Viruses: Retroid viruses: Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.

REFERENCE

AUTHORS Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.  
TITLE Establishment and characterization of molecular clones of porcine  
endogenous retroviruses replicating on human cells  
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)  
MEDLINE 20219394

# REFERENCE

## AUTHORS

Direct Submission

## JOURNAL

Submitted (04-May-1999) Tonjes R.R., Medical Biotechnology,  
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225  
Langen, GERMANY

## FEATURES

### source

Location/Qualifiers

1..8918

/organism="porcine endogenous retrovirus"

/proviral

/strain="type C"

/db\_xref="taxon:61673"

/clone="33"

/note="grown in PK15 (ATCC No. CCL-33) cell line"

1..707

### misc\_feature

545

/note="cap site"

710..727

/note="gly4"

1154..2725

/gene="gag"

1154..2725

/gene="gag"

/function="retroviral capsid"

/product="gag protein"

join(1154..2725,2729..6313,6189..8162)

/note="viral polypeptide"

/codon\_start=1

/protein\_id="CAB65339.1"

/db\_xref="GI:6688946"

/db\_xref="SPTREMBL:Q9Q1X5"

/translation="MGQTVTPTSLTLDHMTVEKSRHNLSVQVKGPMQWTFCASEWP  
TFDVPSEGTFSNELLAVKALIFQSPSSHPDQEPYILTWDLADDPWPKWPNL  
TKRPKPRILALGKNKSAEKVPSPIRYEIEEPTWPEPPVPPPPVPAQAVRG  
PSAPGAPVVEGPAAGTSRRGATPERTDAIILPRTYGPMPGGQLQLOPYWRSS  
ADLYNKTNIHPPFSEDQRLTGLVSEMFSDQDQOQLQTLFTTERERILLER  
RKNVPGADGPTQLQNEIDMGFPLTRFGWDTNABGRESLKTIHQALVAGRGARRP  
TNLAKREVWMPGPNPSPVLEMLERAFRTPTDPTSEAKKASVALFAGSALDIR  
KKLQRLGQLAEALDLVREAEKYIYRRETEKEKEKEKEKEKEKEKEKEKEKE  
KILAAVVEGSSRRERDFRKIRSGPQSGNLNRTLDKQDCAYCKEKHWRNCPK  
KGNKPKVLTEEDKDRGSDPLPEPRVTLKVEGQPVFVDTGABSHVLOPLGLK  
KEKSNWMTAGTGOBYPTWTPTVTLAVGVTHSFLVIEPCVPPLGRDLLTKMGAOI  
SFEQGRPEVSNKPIIVLQLDDEYRLYSPOVKPOODIQSWLEQFPQWAEATAGN  
LAKQVPPVQLKASATPVSRYPQLPSREAREGLWPHQRLIQOGLILVPSPWNTPL  
LPVRKPGTNDYRPVQDLREVNKRVDIHTPTPNPNYLNLSALPPERNMTYVLDLDAFF  
CLRLHPTSQPLAFEPWRDPTGTQTLTQLPQGFKNSTIFDEALHRLDLAFRIQH  
POVTLLOVDDLLLAGATKDCLEGTKALLLELDLGYRASAKAQICRREVTYLGVS  
LRGGORWLTAKKTVQIIPATTAQOVREFELGTAGFCRLWIPIGFATLAAPLYLTK  
KGFSSWAPHEQKAFDAIKKALLSAPALADPTKPTFLYVDERKGVARGVLTQTLGPW  
RRPVYLSKLDLPVAGSWPCLKAIAVAAILVDADKLTIGQNTITVAPHALENIVRQ  
PPDRMTNARMTHYQSLLTERTVFAPPAALNPATLPEETDEPVDHCHOLLEETFG  
VRKDLTIDPLTGEVLTWFTDSSYVVEGKRMAGAAVVDGTHTIWASLPEGTSAQAE  
LMALTOALRLAEGKSIINYDSRYAFATAHVGAITYKORGLTWISSREIKNKEILSL  
LEALHLPKRLAIHCPGHOKAKDLISRGNOMADRVAKQAQAVNLLPIIETPKAPEPR  
QOYTLDEWQEIKKIDQSPSEPTGCTYSYKELIPHKEGLEYYQOIHRLTHLGHKLQ  
OLVTSYTHVLRPGVADSWPCCQLVNPASPTLPPGKRLRGYSGAHWEVDFTF  
VKPAKYKYLLEVDFTSGWVEAYPTKSTVVAKKILEEIPFRFGIPKIVGSDNG  
PAPFAQVSGQLAKILGIDWKILHCAIRPQSSQVERMNTIKETILKTITGINDWTA  
LLPVLFRVNTPGOGLTPEYLLYGGPPLVEIASVHSDALLSQPLSRLLKALEVY  
RORAWKOLREAYSGEGDLOVPRFOVGSVYVRRHRAGNLETWKGPYLLTPTAV  
KVEGISTWHTSHVKLAPPDSDGWRAEKTENPLKLRLHRLVPSYNNNSQOIHTLSW  
RHLTPRGEPKRLRILPISFASIAFWLTLITPQASSKRLDSSNPHPLSLWLLIDP  
DGTVTNSTRGVAPRGTPWPLHFLRLINPAVKSTPNLVRYSYFCCPGTKEKYC  
GDGSEFCRRWSCVTNSDGMKWPISLDQVKFYSFVNSGPKYKVMKLYDKDSCSPD  
LDYLYISFTEKQKQENTKINGMSGIVFYKGGAGSTLRLRIETETPEPVAVG  
GSRKLLTQAGNAINSDPDATSSCHLCLSSGGPPYEGNAKEGKNVYKHNQCTW  
QORVLEAGGPPALPHNLPVOLTSLRPDITOSPSTGLTPTNPRNSPGVPVKT  
PKVLAEGGPPALPHNLPVOLTSLRPDITOSPSTGLTPTNPRNSPGVPVKT  
GRLKLLTEVSGAGTCTIGRAPPSHQHCISTVVEQASENQIILVPGYNRWACNTGL  
TPCVSTVFNQSKDFCVWQIVPRVYHPEVLDYDYNRKRKRPVSLTIAVMIGL  
LGTAVGTGTAAITGQLEKGLHAAETDLRALESVNSLSEVY  
QNRGLDLLFLREGGLCAALKECCCFYVHSGAIRDSMKRLRERLRERERADQGW  
FEGWFRSPWMTLLSALTGLPVLLLLLTGVCPLINRFVAFVRRVRSAYQIMVLRQO  
YQGLLSQGETDL"

```

mat_peptide      2729..6313
/gene="pol"
/function="retroviral protease and polymerase"
gene             2729..6313
/product="Pol protein"
mat_peptide      6189..8162
/gene="pol"
/gene="env"
/function="retroviral envelope"
gene             6189..8162
/product="Env protein"
LTR              8212..8918
/polya_signal    8811..8816
BASE COUNT      2327 a 2245 c 2334 g 2012 t
ORIGIN

alignment_scores:
  Quality: 6232.50      Length: 1195
  Ratio: 5.237          Gaps: 1
  Percent Similarity: 99.582  Percent Identity: 99.079

alignment_block:
US-09-171-553B-5 x PEN133816 ..
Align seg 1/1 to: PEN133816 from: 1 to: 8918

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2729 GGGAGACGGGGTTCGGAGCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 2778
|||||
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2779 GGAGGGCCACCAAGTTCAGTTCCTGGTGTATACCGGAGCGGAGATTTCAG 2828
|||||
34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysLysSerTrpValMet 50
|||||
2829 TGCCTGCTACAAACCATTAGGAAACATAAGAAAGAAAAAATCCTGGGTGATG 2878
|||||
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgThrValas 67
|||||
2879 GGTGCCACAGGGCAACGGCAGTATCCATGGACTACCCGGAAGAACCGTTGA 2928
|||||
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
2929 CTGGCAGTGGAGCGGTAAACCCACTGTTCTGGTCTATCCCTGAGTGCC 2978
|||||
84 roValProLeuLeuGlyArgAspLeuThrLysMetGlyAlaGlnIle 100
|||||
2979 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 3028
|||||
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
3029 TCTTTTGAACAAGGAAGAACAGAGTCTCTGTGAATAACAACCCATCAC 3078
|||||
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
3079 TGTGTTGACCTTCAAAATAGATGATGATGATGATGATGATGATGATGATG 3128
|||||
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
3129 TAAAGCCTGATCAAGATATACAGTCTCTGGTTGGAGCAGTTTCCCAAGCC 3178
|||||
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
3179 TGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCCTCCACAGGT 3228
|||||
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
3229 TATTCACTGAAGCCAGTGTCTACACAGTATCAGTACAGTACAGTACCCCT 3278
|||||
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||

```

```

|||||
3279 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCCATGTTCAAAGATTAAATC 3328
|||||
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
3329 CAACAGGGCATCCTAGTCTCTGTCCTCAATCCCTTGAATACATCCCTGCT 3378
|||||
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
3379 ACCGGTTAGGAAGCCTGGGACCAATGATTATCGACCAGTACAGGACTTGA 3428
|||||
234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
3429 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCT 3478
|||||
251 TyrAsnLeuLeuSerAlaLeuProGluProGluArgAsnTrpThrValle 267
|||||
3479 TATAACCTCTTGAGCGCCCTCCGCTGAACGGAACCTGCTACACAGTATT 3528
|||||
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
3529 GGACTTAAAGATGCCCTCTCTGCTGCTGAGATTACACCCACTAGCCAC 3578
|||||
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3579 CGCTTTTTCGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3628
|||||
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3629 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTTT 3678
|||||
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3679 TGACGAAGCCCTACACAGGGACCTGGCCAACTTCAGGATCCCAACACCTC 3728
|||||
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3729 AGGTGACCTCTCCCTCAGTACGTGGATGACCTGCTTGGCGGGAGCCACC 3778
|||||
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367
|||||
3779 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTTGCTGA 3828
|||||
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3829 CCTAGCTACAGAGCTCTGCTAAGAGGCCAGATTTGCAGGAGAGAGG 3878
|||||
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3879 TAACATACCTGGGGTACAGTTTTCGGGGCGGGCAGCGATGGCTGACGGAG 3928
|||||
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
|||||
3929 GCACGGAAGAAACTGTAGTCAGATACCGGCCCAACCAACACAGCCAAACA 3978
|||||
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3979 AGTGAGAGAGTATTTGGGCACAGCTGGATTTTGCAGACTGTGTGATCCCGG 4028
|||||
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlyLysGly 450
|||||
4029 GGTTCGCGACCTTAGCAGCCCTCTACCCGCTAACCAAAAGAAAGGG 4078
|||||
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
4079 GAATTCCTCTGGGCTCCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 4128
|||||
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
4129 GGCCCTGCTGAGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4178
|||||
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
|||||

```

4179 TTACCTTTATATGTGGATGAGCGTAAGGAGTAGCCGAGGAGTGTAAACC 4228  
501 GlnThrLeuGlyProTrpArgProValAlaThrLeuSerLysLe 517  
4229 CAAACCTTAGGACCATGGAGGAGCTGTTCCTACCTTGTCAAGAAGCT 4278  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaLeAlav 534  
4279 TGATCCTGTAGCAGTGTGTGGCCGCTATGCTGAAGGCTATCGCAGCTG 4328  
534 AlaAlaLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
4329 TGGCCATACCTGGTCAAGACGCTGCAAAATTCACCTTTCGGACAGAAATA 4378  
551 ThrValIleAlaProHisAlaLeuGlnAsnIleValArgGlnProProAs 567  
4379 ACTGTAAATAGCCCCCATGTGATGGAGAACATCGTTTCGGACAGCCCCAGA 4428  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
4429 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCGTGTCTCA 4478  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
4479 CAGAGAGGCTCACTTTCGCTCCACGAGCGCTCTCAACCCCTGCCACTCTT 4528  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
4529 CTGCTGAAGAGACTGATGAACCACTGACTCATGATGCCATCAACTATT 4578  
617 uLeuGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4579 GATTGAGGAGAGTGGGTCGCGAGGACCTTACACATACCGCTGACTG 4628  
634 lYgluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
4629 GAGAAGTGTCACTACCTGGTTCACCTGACGGAAGCAGCTATGTTGGTGAAGGT 4678  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4679 AAGAGGATGGCTGGGCGCGCGTGTGGACGGGACCCACACGATCTGGGC 4728  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4729 CAGCAGCCTGGCGGAAGAACTTACGCGCAAAAGCTCAGCTCATGGCCC 4778  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4779 TCACCAAGCTTTCGGGCTGGCCGGAAGGAAATCCATAAACATTTATACG 4828  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
4829 GACAGCAGGTATGCTTTCGCACTGCACACGCTACACGGGGGCATCTATAA 4878  
717 scInArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4879 ACAAGAGGGGTTCCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4928  
734 lUileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4929 AAATTCCTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4978  
751 lIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4979 ATACACTGTCTTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 5028  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
5029 CCAGATGGCTGACCGGGTGGCAAGCAGCGCCCGGCTGTTAACCTTC 5078  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnThrThr 800  
5079 TGCCTATAATAGAACGCCCAAGGCCCAAGCCCAAGACCCAGACAGCTACACC 5128

801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
5129 CTAGAGACTGGCAAGAGATAAAAGATAGACCACTTCTCTGAGACTCC 5178  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
5179 GGAGGGGACCTGCTATACCTCATATGGAGAGAAATCCTGCCACAAAAG 5228  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
5229 AAGGGTTAGAAATATGTCACAGATACCTGCTCACCCACCTAGGAAC 5278  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
5279 AAACACCTGCAGCAGTTGGTCAGAACATCCCCCTATCATGTTCTGAGGCT 5328  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
5329 ACCAGGAGTGGCTGACTCGGTGTCAAACATTTGTCCTGCCCTGCCAGCTG 5378  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
5379 TTAATGCTAATCCTTCCAGAACTACCTCCAGAAAGAGACTAAGGGGAGC 5428  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
5429 CACCCAGGCGCTCCTGGAAGTGGACTTCACCTAGGTAAAGCCGCTAA 5478  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
5479 ATACGGAACAAATATCTATTTGTTTGTAGACACCTTTTTCAGGATGG 5528  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
5529 TAGAGGCTTATCCTACTAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 5578  
951 lLeuLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
5579 ATACTGGAGGAAATTTTCGAGATTTGGAATACCTAAGGTAAATCGGGTC 5628  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5629 AGCAATATGGTCCAGCTTTTGTCCAGGTAAGTCAGGACTGGCCAGA 5678  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5679 TATTGGGATTTGATTGSAATCTCATTTGTCATACAGACCCCAAGCTCA 5728  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
5729 GGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGACCCCTTACCAAT 5778  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuPheValL 1034  
5779 GACCACAGACTGGCAATTAATGATGGATAGCTCTCTGCTCCCTTTGTGC 5828  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
5829 TTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5878  
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067  
5879 TTGCTCTACGGGGGACCCCGGTTGGTAGAAATTCCTCTACATAG 5928  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
5929 TGCTGATGCTGCTGCTTCCAGGCTCTGTCTCTAGGCTCAAGGCGCTCG 5978  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
5979 AGTGGGTGAGGCAACGAGCGTGGAAAGCAGCTCCGGGAGGCGCTACTCAGGA 6028

```

1101 ...GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValty 1116
|||||
6029 GAAGGAGACTTGCAGTTCACATCGCTTCCCAAGTGGAGATTTCAGTCTA 6078
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
6079 TGTTAGACGCCACCGTGCAGAAACCTCGAGACTCGGTGGGAAGGCCCTT 6128
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
6129 ATCTCGTACTTTCACACACCAACGCTGTGAAGTCGAAGAAATCTCC 6178
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166
|||||
6179 ACCTGGATCCATACATCCAGCTTAAGCTGGCGCACCTCCCGACTCGGG 6228
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6229 GTGAGAGCGGAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCC 6278
1183 euValProTySerAsnAsnSerProGlyGln 1194
|||||
6279 TGGTTCCTTACTTAACAATAACTCCCGACGCCAG 6313
seq_name: gb_vi:PEN293657

seq_documentation_block:
LOCUS PEN293657 8763 bp DNA VRL 30-MAY-2001
DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene
and env gene, class B, clone 213.
ACCESSION AJ293657
VERSION 1
KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;
polymerase; polypeptide.
SOURCE Porcine endogenous type C retrovirus.
ORGANISM Porcine endogenous type C retrovirus.
REFERENCE 1 (bases 1 to 8763)
AUTHORS Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
TITLE Krach, U., Fischer, N., Czaderna, F. and Tonjes, R. R.
Comparison of replication-competent molecular clones of porcine
endogenous retrovirus class A and class B derived from pig and
human cells
J. Virol. 75 (12), 5465-5472 (2001)
MEDLINE 21256017
REFERENCE 2 (bases 1 to 8763)
AUTHORS Tonjes, R. R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2000) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,
GERMANY
FEATURES
source Location/Qualifiers
1..8763
/organism="porcine endogenous type C retrovirus"
/proviral
/specific_host="Sus scrofa"
/db_xref="taxon:70540"
/clone="213"
/cell_lines="PK15"
/note="class B"
1..630
1..467
LTR misc_feature
primer_bind /note="Cap site"
gene 633..650
1077..2651
/gene="gag"
1077..2651
/gene="gag"
/gene="gag"
/codon_start=1
/product="group specific antigen"
/protein_id="CAC39625.1"
/db_xref="GI:14275842"
/trans_lation="MGQTVTTPLSLTLDHWTVEKRAHNLVSQVKGKGPWOTFCSEWP
TFDVGWPEGTFSNELLAVKAIIFQTGPGSHPDQEPYILTQDLAEEPPPPWKPWLN

```

```

KPRKPGRIILALGKHNKSAEKPSPGIYPIETEEPTWPEOPVPPPPYPAQAVRG
PSAPGAPVYEGPAAGT"RSRGATP"PERTDELAIALPLRTYGPMPGQGLQPLQWPFSS
ADLYNWKTNHPFSDFOELMTGLVLESIMF"SHQPTWDDCQLLOT"LFTEERERILLLEA
RKNVFGADGRPTOLQNEIDMGITRPGWYDNTAEGRESLK1YRQALVAGLRGASRRP
TNLAKVREVMQGPNEPSPVFLERLMEAFRRFTFPDPTSEAKASVALAFTCQSALDIR
KKLQRLGELQAEALDLVREAKEYYRRETTEEKEQKEREERERRRORRQENKT
KILAAVVEGKSSRRERDRFRKIRSGPQSGNLGNRTPLDKDQCAYPEKCKEKGHWARNCPK
KNGKPKVLALDEEDK"
2652..6239
/gene="pol"
<2652..6239
/feature="pol"
/note="complete pol ORF, gag-pol is expressed as a
polyprotein"
/citation=[1]
/codon_start=1
/product="polymerase"
/protein_id="CAC39626.1"
/db_xref="GI:14275843"
/trans_lation="GRRGSDPLPEPRVTLKVEGQPVVEFLVDTGAHSHVLLQPLGKLKE
KSNWMTGATGQRPWTTRTVDLAGRVTHSELVIPCVPVLLGRDLITKMGQISF
EQGRPEVSVANNKPITVLTQLDDEYRLYSPQVKPDQDIOSMLEQFPQAWAETAGMLA
KQVPPQVQLKASATPVSVROYPLSREAREGIMWPHQRLIQOGLIPVSPWMTPLLP
VRKPGTNDYRPVODLREVNRKRVODIHTPTNPYNLLSALPERNWTVLIDLKDAFCL
RLHPTSOPLFAFEWRDPTGRTGOLTWRLPQGNKSPTEFEALRDLANFRIHQHQ
VTLLQYVDDLLLAGATKQDCLESTKALLLESLGLYRASAKKAQICRRVYILGYSLR
GQWMLTEARKKTVYQIPATAKQVREFLTAGFCRLMIPGFATLAAPLYPLTKBKG
EFWMAHQKAFDAIKALLSAPALAPDVTKPTFLYVDERKGVARGVLTQTLGPPWR
PVALSKLDPVAGSNPVCILAAVALILVKDADKLTLGONITVIAHALENIRQPP
DRWNTARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTBCHOLLIEBTVPR
KDLTDIPLTGEVLTVTDSGSSVVEGKMAAGVDTGRTIWAASSIPEGTSQAKELM
ALQALRLAEGKSNITVDSRYAFATAHVGAIGYKQGLITSAGREIKKKEELSLLE
ALHPLKRALIHCPCGOKADLISRGONMADRVAQAQAANLLPIETPAKPRRQ
YTLEDWQEIKKIDOFSEPTCYTSYKELPHKEGLEVQOIHRLTLHGLTKHLQOL
VRTSPYHLRLPGVADSVKHCVPQILVNPANRIPPRKRLRGSHCAHWEVDTEVK
PAKYGNKYLIVVDFTSGWVEAYPTKSTVAVKILIEIETPRFGIPKIVGSDNGPA
FVAQVSGOLAKILGIDMKLHCAVPOSSGOVERMNTIKETLTKLTETRGINDWALL
PFVLFVRNTPGOGLTTPYELLYGGPPPLVEIASVHASADVLLSQPLSRKLALEWRQ
RANKQLEAYSGGEDILQPHRFQVGSVYVRHAGNLTRMKGPVYLVLTTPTTAVKV
EGISTWIHASHVKLAPPPDSGWAERKTENPLKRLRLHLRPVYSNNNSPGQ"
6112..8085
/gene="env"
6112..8085
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/trans_lation="MHPTLSWRHLPTRGGEKPLRITPLSFASWTAWPLTTLITPQASSK
RLIDSNPHRPLSLTLWLIIDPDGVTNVSTRGVAPRGTWHPHELHCLRLINFAVKSSTP
PNLVRSGIFCCPGTEKEKTCGGSGESFCRRSCVTSNDGDKWKPISLQDRVKFSFVN
SGPGYKVMKLYKDKSCPSDLDLYKISFTEKQKQENIQKINGMSWGLVFKYGGGA
GSTLTIRLETTEPPVAVPKDLAEQGPAPLEPHNLPPVQLTSLRPDITQPPSX
STTGLTINTPRNSGPVVKTGORLFSLIQGAFOALNSTDPDATSCWCLCSGGPPY
EGMAKEGKENVTKHNRNCTWGSRNKLTLEYSKGTCIGKAPPSHOHLNSTVYVEQ
ASENOYLVPQYNHWACNTGLPCVSTSVNOSKQDCVAVQIVPRVYHPEBWLDEY
DYRNPDKREPSTLTLAVMLGTAVGVTGTAALITGQQLEKGLGEUHAAMTEDLR
ALEESVNLSESLTSEVVLQNRRLGLDLFLREGGLCAALKECCFYVDHSGAIRDS
MSKLRELRERRRRRERDQGFEGFNFRSPWMTTLTALSALTGPLVLLLTTLTGCPCLN
RFVAFVRERVSAVQIMVLRRQYQGLLSQGETDL"
8133..8763
8655..8660
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
LTR
polyA_signal
8655..8660
1 others
alignment_scores:
Quality: 6229.50 Length: 1195
Ratio: 5.235 Gaps: 1
Percent Similarity: 99.582 Percent Identity: 98.996
alignment_block:
US-09-171-553B-5 x PEN293657

```

Align seg 1/1 to: PEN293657 from: 1 to: 8763

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
2652 GGGAGACGGGGTTCGGAGCCCTCCCGAGCCGAGGGTAACTTTGAAGGT 2701  
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
2702 GGAGGGCAACACAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTCAG 2751  
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
2752 TGCTGCTACACCATTAGGAACTTAAAGAAATAAATAATCCTGGGTGATG 2801  
51 GlyAlaThrGlyGlnArgGlnProThrArgArgThrValAs 67  
2802 GGTGCCACAGGCAACGGCAGTATCCATGGACTACCCGAAACCGTTGA 2851  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValileProGluCysP 84  
2852 CTTGGCAGTGGGACGGGTAACCCACTCGTTTCTGGTGCATCCTGAGTGCC 2901  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnile 100  
2902 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2951  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
2952 TCTTTTGAACAGGAGGACCAAGTCTGTGTAATAACAACCCATCAC 3001  
117 rValLeuThrLeuGlnLeuAspGluThrArgLeuTyrSerProGlnV 134  
3002 TGTGTTGACCCCTCCAAATTAGATGATGAATATCGACTATATCTCCCCAA 3051  
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
3052 TAAAGCCTCATCAAGATATACAGTCTCTGTTGGAGCAGTTTCCCCAAGCC 3101  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
3102 TGGGCAGAAACCGCAGGGATGGTTTGGCAAGCAAGTTCCCCACAGGT 3151  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
3152 TATTCAACTGAAGGCCAGTGTACACAGTATCAGTCAGACAGTACCCCT 3201  
184 euSerArgGluAlaArgGluGlyIleTyrProHisValGlnArgLeuile 200  
3202 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGATGTTCAAAGATTAATC 3251  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
3252 CAACAGGCATCCTAGTTCTCTGCCAATCCCTTGGAAATCTCCCTGCT 3301  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
3302 ACCGGTTAGGAAGCCTGGGACCAATGATATACGACGAGTACAGGACTTGA 3351  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
3352 GAGAGTCAATAAAGGGTCGAGGACATACACCCCAACGGTCCCGAACCCCT 3401  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
3402 TATAACCTCTTGAGCGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 3451  
267 uaspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
3452 GGACTTAAAGATGCTTCTTCTGCTGAGATTTACACCCCTAGCCCAAC 3501  
284 roLeuPheAlaPheGluThrArgAspProGlyThrGlyArgThrGlyGln 300

3502 CACTTTTTCCTTCGAAATGGAGAGATCCAGGTACGGGAAGAACCCGGCAG 3551  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
3552 CTCACCTGGACCCGACTGCCCAAGGTTTCAAGAACTCCCGACATCTT 3601  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3602 TGACGAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAAACACCCTC 3651  
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
3652 AGGTGACCCCTCCAGTACGTGATGACCTGCTTCTGGGGAGGCCACC 3701  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
3702 AAACAGGACTGCTTAGAAGGTACGAAGGCACACTACTGCTGGAATGTCTGA 3751  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3752 CCTAGGCTACAGAGCCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGG 3801  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
3802 TAACATACCTTGGGTACAGTTTGGGGCGGCGCAGCGATGGCTGACGGAG 3851  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
3852 GCACGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACAGCCAAACA 3901  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
3902 AGTGAGAGAGTTTGGGGACAGCTGGATTTTCAGACTGTGGATCCCGG 3951  
434 lyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLysGly 450  
3952 GGTTTGCGACCTTAGCAGCCCTCTACCCGCTAACCAAGAAAGG 4001  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
4002 GAATTCCTCTGGGCTCTGAGCACACAGAGGCATTTGATGCTATCAAAA 4051  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
4052 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCT 4101  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
4102 TTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGAGGAGTTTAAACC 4151  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
4152 CAAACCTTAGGACCATGGAGAGACCTGTGCTTACCTGTCAAGAAAGCT 4201  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaVal 534  
4202 TGATCTGTAGCCAGTGGTTGGCCCGCTGATGCTGAAGGCTATCGCAGCTG 4251  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
4252 TGGCCATACTGCTCAAGGACGCTGACAAATGACTTTGGGACAGAAATATA 4301  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567  
4302 ACTGTAATAGCCCCCATGCAATTGGAGAACATCGTTCGGCAGCCCCCAGA 4351  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
4352 CCGATGATGACCAACGCCGCCATGACCCACTATCAAGGCTGCTTCTCA 4401  
584 hrGluArgValThrPheAlaProProAlaLaLeuLeuAsnProAlaThrLeu 600  
4402 CAGAGAGGCTCACITTCGCTCCACAGCGCTCTCAACCCCTGCCACTCTT 4451

601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
4452 CTGCCTCAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 4501  
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProIleuThrg 634  
|||||  
4502 GATTGAGGACACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4551  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4552 GAGAAGTGCCTAACCTGGTTCCACTACGCGAAGCAGCTATGTGGTGGAGGT 4601  
651 LysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4602 AAGAAGATGGCTGGAGCGGGTGGTGGACGGGACCGCACGATCTGGGC 4651  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
4652 CACAGCCTGCCGGAAGGAACCTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4701  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4702 TCACGCAAGCTTTGGCGCTGGCCGAAGGGAATCCATAACATTTATACG 4751  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrIly 717  
|||||  
4752 GACAGCAGGTATGCCCTTTGGGACTGCACACAGCTACACGGGGCCATCTATAA 4801  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4802 ACAAGGGGGTTGCTTACCTCAGCAGGGAGGGAATAAAGAACAAAGAGG 4851  
734 luileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4852 AAATTTCAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4901  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
4902 ATACACTGCTCGACATCAGAAAGCCAAAGATCTCATATCTAGAGGGAA 4951  
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784  
|||||  
4952 CCAGATGGCTGACCGGGTTGCCAAGCAGCGAGCCCGAGGCTGTAAAGCTTC 5001  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnIlyThr 800  
|||||  
5002 TGCTTATAATAGAAACGCCCAAGCCCAAGACCCAGACAGCAGTACACC 5051  
801 LeuGluAspTrpGlnIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5052 CTAGAAGACTGGCAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCC 5101  
817 oGluGlyThrCysTyrThrSerTyrClyLysGluIleLeuProHisLysG 834  
|||||  
5102 GGAGGGACCTGCTATACCTCATATGGGAAGGAATCTTCCGCCCAACAAAG 5151  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
5152 AAGGGTTAGAAATATGTCACACAGATACATCGTCTCACCCACCTAGGAAT 5201  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
5202 AAACACCTGCAGCAGTTGGTCTAGAAACATCCCTTATCATGTTCTGAGGCT 5251  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
5252 ACCAGAGTGGCTGACTCGGTGGTCAACATTTGTGGCCCTGCCAGCTGG 5301  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
5302 TTAATGCTAATCTTCCAGAAATACCTCCAGGAAAGAGAGACTAAGGGGAGC 5351

seq\_name: gb\_vi: PEN133817

901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLys 917  
|||||  
5352 CACCCAGGCGCTCACTGGGAAGTGGAGCTTCACTGAGGTAAAGCCGGCTAA 5401  
917 StyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
5402 ATACGGAACAAATATCTATTTGGTTTTGTAGACACCTTTTCAGATGGG 5451  
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
5452 TAGAGCTTATCTCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAAAA 5501  
951 lleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967  
|||||  
5502 ATACTGGAGAAATTTTCCGAGATTGGGAATACCTAAGGTAATCGGGTC 5551  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5552 AGACAATGGTCCAGCTTTGTTGGCCAGGTAAAGTCAGGAGCTGGCCAAGA 5601  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
5602 TATTGGGATTTGATTTGGAAACTGTCATTGTGCATACAGACCCCAAGCTCA 5651  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
5652 GCACAGGTAGAGAGGATGAATAGAACCATTAAGAGAGACCTTACCAAAAT 5701  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
5702 GACCACAGAGACTGGCATTAATGATTGGATAGCTCTCCTGCCCTTTGTGC 5751  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
5752 TTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGCTGACCCCTATGAA 5801  
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5802 TTGCTCTACGGGGGACCCCGCTTGGTAGAAATTTGCTTCTGTACATAG 5851  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
5852 TGCTGATGTGCTGCTTCCAGCCTTTGTTCTTAGGCTCAAGCGGCTCG 5901  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
|||||  
5902 AGTGGGTGAGCAACGACGCTGGAGAGCAGCTCCGGGAGGCTACTCAGGA 5951  
1101 ... GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116  
|||||  
5952 GAAGGAGACTTGCAGATTCCACATCGCTTCCAAGTGGGAGATTTCAGTCTA 6001  
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133  
|||||  
6002 TGTTAGACGCCCATCGCAGGAAACCTCGAGACTCGGTGGGAAGGGCCCTT 6051  
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
|||||  
6052 ATCTCGTACTTTTGCACACACCAACGCTGTGAAAGTGAAGGAATTCCTC 6101  
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGly 1166  
|||||  
6102 ACCTGGATCCATGCTCCACGTTTAAAGTGGCGCCACCTCCCGACTCGGG 6151  
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
|||||  
6152 GTGGAGAGCCGAAAGAGCTGAGAAATCCCTTTAAGCTTTCGCTTCATGCC 6201  
1183 euValProTyrSerAsnAsnSerProGlyGln 1194  
|||||  
6202 TGGTTCCTTACTCTTAACAAATTAACCTCCCGAGGCCAG 6236

```

seq_documentation_block:
LOCUS      PEN133817      8849 bp      DNA              VRL              07-JAN-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
            genes and LTR (Class A, clone 42).
ACCESSION  AJ133817
VERSION    1
KEYWORDS   env gene; env protein; gag gene; gag protein; pol gene; pol
            protein.
SOURCE     porcine endogenous retrovirus.
ORGANISM   Porcine endogenous retrovirus
            Viruses; Retroviral viruses; Retroviridae; Mammalian type C
            retroviruses; 1-Mammalian type C virus group.
REFERENCE  1 (bases 1 to 8849)
            Czauderna, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and
            Toenjes, R.R.
            Molecular Characterization of Human-tropic and
            Replication-competent Porcine Endogenous Retroviruses
            Unpublished
            2 (bases 1 to 8849)
            Toenjes, R.R.
            Direct Submission
            Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
            Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
            Langen, GERMANY
FEATURES             Location/Qualifiers
     source           1..8849
                     /organism="porcine endogenous retrovirus"
                     /proviral
                     /strain="type C"
                     /db_xref="taxon:61673"
                     /clone="42"
                     /note="grown in PK15 (ATCC No. CCL-33) cell line"
     LTR              1..668
     misc_feature     506
                     /note="cap site"
     primer_bind      671..688
                     /note="gly4"
     gene             1115..2686
                     /gene="gag"
                     1115..2686
                     /gene="gag"
                     1115..2686
                     /gene="gag"
                     /function="retroviral capsid"
                     /product="gag protein"
                     join(1115..2686,2690..6271,6150..8132)
                     /note="viral polypeptide"
                     /codon_start=1
                     /protein_id="CAB65340.1"
                     /db_xref="GI:6688948"
                     /translation="MGQTVVTPLSLTDHWTVEVRSRAHNLVQVKKGPWQTCASEWP
TFDVGWSEGTFSNSELILAVKAIIFQFGSHPDQEPYILTWDLAEDPPWVKPWLN
KPRKPGPRILALGKNKHSKAEKVEPSRIYETEEPTWPEQPPVPPYPYPAQAVRG
PSAPGAPVVEGPAAGTRSRGATPERTDEIALPLRTYGPMPGQLOPLOWPFS
ADLYNKNTHNPPESEDPQRLTGLVESLIMFSHOPTWDDCOQLQTLFTTEERILLLEA
RKNVPGADGRTQLONEIDMGFFLTPRGWDYNTAEGRESKLIYRQALVAGLGRARRP
TNLAKREVWQGNPEPSFLELRMEAFRRFTPDPTSEAQKASVALATIGSALDIR
KILQRLGLEQAEALDLVRAEAYVYRRETEERKEQREEREERDRQKRLNT
KILAAVEGKSSRERDRFRKIRSGPRQSGNLGNRTPLDKDQCAKCKEKGHWARNCPK
KGNKGPVLALEDKDGRRGSDPLPEPRVTLKVEGQPVREFLVDVTAHESVLLQPLGL
KKNKSWVMGATGQRYPTTRTRVTLVGVGVTHSVFLVPECPVPLGRDLLTKMGAQI
KEKKSVEVSVNNKPTIVTLQLDDDEIRIYSPQVKDDQIQSLQWLEFPQAWAETAGM
SFQGRPEVSVNNKPTIVTLQLDDDEIRIYSPQVKDDQIQSLQWLEFPQAWAETAGM
LAKOVQVITQLKASATPVSVRQPLSREAREGINWPHVORLIQGLILVYVQSPWNPPL
LPVKRPTNDYREVQDLREYNKRVQDIHTVNPVNLKLSALPERNMYTVLDLKDAFF
CLRLHPTSDVLFQFVWRDPTGTGTGTWTRLPQGEKNSPTIFDEALHRLDLANFKIQH
POVTLQYVDDLALAGATKODCLEGTAKLLESLDLGYSRASKAIOICRREYLYIGYS
LRGQWPLTEARKKTVQYIPATTAQVREFLGTAGFCRLWIPGFATLAAPLYPLTKE
KGFSWAPENQKAFDAIKKALLSAPALPDVTKPTLYVDERKGVARGVLTQTLGPV
RRPVVAUSKKLDVPSVGVNCLKAAVALVLDKADKLTGLQNTIVIAHALLENVIRQ
PPDRMTNMTNHYOSLLTERFTVFPAPALNPATLPEETDEPETHDCHQLLIETG
VRKDLTDIPITGVLTWFTDCSSYVVEGKMAAGVVDGTRTITWASSLPEGSAKAE
LMALTOALRAEGKSYNIYDSRYAFATAHVGAIVYKQRLGTLTISAGREKKNKEILSL
LEALHPLKRLAITHCPCHKAKDLISRGNMADRAVAKQAQVNLPIITETPKAPPR
ROYTLEMDQEIKKIDQFSETPGTCYTSYKELLPLPHKEGLEVYVQIHLRHLTLGKHLQ

```

```

OLVRTSPYHVLRPGVADSVVVKHCVPCQLVNANPSRIIPGKRLRGSHGHAHEVDFIE
VKPAQYGNKYLLVDFVDFSGVWEAYPTKKTSTVYAKKILLEEIPFRGIPKVGSDNG
PAFVAQVSGQAKILGIDWKLHCAYRQSSQGVMMRTIKETLTKLTETGINDMMA
LLPVLFRVNTPGOGLTPVELLYGGPPLVEIASVHADVLLSQPLFSRLKALEWV
RQRAKQOLREAYSGEGDQVRFQVGDVSVVRHRAGNLETRWKQPYLVLLTPTAV
KVERIPTWIHSHVAKPAPPPDSGKAKETNPCLKRLHRVYVPSYNNSSSMHPTLSRR
HLPIRGKPKRLKIPLSFASIAWELTSLITPOVNGKRLVDSPNHKLPLSLTLLTDSG
TGININSTQGEAPLGTWMPELYVCLRSVIFGLNQDAPDPDLYRAYFYVCPGPNNEE
QYGNPQDFCFKOWSCITSDNGNMKWPVSQDRVSYSFNNTPSYNYGHRWKMDQ
QVOKDVNRKQISCHSLDLDYKISFTEKQENIKQVINGISWGIYVYGGSGRKGGS
VILIRLETOMEPPEVIAIGNKLAEGOPPIOEQORPSPNSDYNTTSGSVTEPNITI
KITAKLFSLIQCAQALNSTTPEATSSCHLCLASGPYPYEGMARGKENVYKHEKDOC
TWGSONKLTITFEVSGKGCICWPPSHQHLNHTAFNRSESYLYLPGVDVRRWACNT
GLTFCVSTLVFNQKDFCMVQIVPRVYIYVEKAVLDKIDYIRNRPREPISLFLAVM
LGLVAGACGTGTAALITGPQOLEKSLNLRIVTEDLOALEKSVSNLESLSLSEAO
VLQNRGLDLFLKREGICVAKKECCFYVDHSGAIRDMSKLRERLERREREADQ
GWEGFNRSPWMTTLLSALTGPLVLLLLLTPLVCLINREAVFVRERSAVQIMVLR
QQYQGLLSQGETDL"
2690..6271
/gene="pol"
/function="retroviral protease and polymerase"
/product="Pol protein"
2690..6271
/gene="pol"
6150..8129
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6150..8129
/gene="env"
8182..8849
8742..8747
polyA_signal
BASE COUNT 2355 a 2191 c 2281 g 2022 t
ORIGIN

alignment_scores:
Quality: 6225.50      Length: 1192
Ratio: 5.245          Gaps: 1
Percent Similarity: 99.581 Percent Identity: 99.161

alignment_block:
US-09-171-553b-5 x PEN133817

Align seg 1/1 to: PEN133817 from: 1 to: 8849

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2690 GGGAGACGGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACTTTGAAGGT 2739

17 LGLuGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2740 GGAGGGCAACCACTGAGTTCCTGTTGATACCGGAGCGGAGCATTCAG 2789

34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysSerTrpValMet 50
|||||
2790 TGTGCTACAAACCATAGGAAATAAAGAAAAAATCCTGGGTGATG 2839

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
|||||
2840 GTGCCACAGGGCAACGCGAGTACCATGACATCCCGAAGAACCGTTGA 2889

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
2890 CTTGGGAGTGGGAGGGGTAAACCCACTCGTTCTGTCATCCCTGAGTGCC 2939

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGluIle 100
|||||
2940 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAATT 2989

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2990 TCTTTTGAACAAGGAGACCAAGAGTCTCTGTGAATAACAACCAACATCAC 3039

```



117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
|||||  
3040 TGTGTGACCTCCAAATAGATGATGAATATCGACTATATCTCCCCAAG 3089  
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
|||||  
3090 TAAAGCCTGATCAAGATATACAGTCCCTGGTTGGAGCAGATTTCCCAAGCC 3139  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
|||||  
3140 TGGCGAAGAACCGCAGGATGGGTTTGGCAAGCAAGTTCCTCCACAGGT 3189  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
|||||  
3190 TATTCAACTGAAGGCCAGTCTACACAGATATCAGTCAGACAGTACCCCT 3239  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
|||||  
3240 TGAGTAGAGAGGCTCGAAGAGGATTTGGCCGCATGTTCAAGATTAAATC 3289  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
|||||  
3290 CAACAGGGCATCTAGTTCCTGTCCTCCCAATCCCTTGGAAATACTCCCTGCT 3339  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
3340 ACCGGTTAGGAAGCCTGGGACCAATGATTATCGACCAGTACAGACTTGA 3389  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
3390 GAGAGGTCAATAAAGGGTGCAGGACATACACCAAGCGTCCCGAACCTT 3439  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
|||||  
3440 TATAACCTCTTGAGCGCCCTCCCGCTTGAACGGAACTGGTACACAGTATT 3489  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
|||||  
3490 GGACTTAAAGATGCCCTTCTCTGCTGAGATTACACCCCACTAGCCAAC 3539  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3540 CACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3589  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3590 CTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCCCGACCATCTT 3639  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3640 TGACGAAGCCCTACACAGGACCTGGCCAACCTTCAAGATCCAAACACCCCTC 3689  
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
|||||  
3690 AGGTGACCTCTCCAGTAGTGGATGACCTGCTTCTGGCGGAGCCACC 3739  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
|||||  
3740 AAACAGGACTGCTTACAGGTACGAAGGCCTACTCTCTGGAATTTGCTGA 3789  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
|||||  
3790 CCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAGG 3839  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
|||||  
3840 TAACATACTTGGGGTACAGTTTCCGGGGGGGGCAGCGATGCTGACGGAG 3889  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
|||||  
3890 GCACGAAGAAACTGTAGTCCAGATACCGGGCCCCAACACAGCCCAACA 3939  
717 sglnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
3940 AGTGACAGAGATTTTGGGACAGCTGGATTTTGCAGACTGTGTGATCCCCG 3989  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
3990 GGTTCGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAGGAGG 4039  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
4040 GAATTCCTCTGGGCTCTCAGCACCAGAAAGGCAATTTGATGCTATCAAAA 4089  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
4090 GGCCTCTGCTGAGGCACTGCTCTGGCCCTCCCTGACGTAACTAAACCTT 4139  
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500  
|||||  
4140 TTACCCCTTTATGTTGATGAGCGTAGGGAGTAGCCCGAGGAGTTTAAACC 4189  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
4190 CAACCCCTAGGACCATGGAGGAGACTGTTGCTTACCTGTCAAGAAAGCT 4239  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534  
|||||  
4240 TGATCCTGTAGCCAGTGGTTGCCCGTATGCTGAAGGCTATCGCAGGTG 4289  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
4290 TGGCCATACTGTCAGGACGCTGACAAATGACTTTGGGACAGAAATATA 4339  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
4340 ACTGTAATAGCCCCCATGCAATTGGAGAACATCTGTTGGCAGCCCCCAGA 4389  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
|||||  
4390 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGACCTGCTTCTCA 4439  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
4440 CAGAGAGGTCACTTTGCTCTCCACGACGCGCTCTCAACCCCTGCCACTCTT 4489  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
4490 CTGCTCTGAAGAGACTGATGAACAGTACTCATGATTGCCATCAACTATT 4539  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4540 GATTGAGGAGACTGGGTCGCCCAAGGACCTTACAGACATACCGCTGACTG 4589  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4590 GAGAAGTGTAACTTGGTTCACTACGGAAGCAGCTATGTTGTTGGAAGGT 4639  
651 LysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4640 AAGAGGATGGCTGGGGCGGCACTGGTGGACGGGCCCGCACGATCTGGGC 4689  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
4690 CAGCAGCCTGCCGGAAGCACTTCAGCCCAAGGCTGAGCTCATGGGCC 4739  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4740 TCACCAAGCTTTGGGCTGGCCGAAGGAAATCCATAACATTTATACG 4789  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4790 GACAGCAGGTATGCTTTTGGGACTGCACACGTACACGGGGCCATCTATAA 4839  
717 sglnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734

```

|||||
4840 ACAAGGGGCTTGCCTTACCTCAGCAGGGGAGGAAATAAGAACAAAGAGG 4889
734 luileuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
4890 AAATCTTAAGCCTATTAGAGCCTTACATTTGCCAAAAGGCTAGCTATT 4939
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
4940 ATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4989
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAlaLeuL 784
4990 CCAGATGGCTGACCGGTTGCCAAGCAGCGACGCCAGGCTGTAAACCTTC 5039
784 euProIleLeuGluThrProLysAlaProGluProArgArgGlnThr 800
5040 TGCCATAATAGAAAGCCCAAGGCCCAAGACCCAGACGAGTACAC 5089
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
5090 CTAGAGACTGCGCAGAGATATAAAGATAGACCACTTCTCTGAGACTCC 5139
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
5140 GGAGGGGACCTGCTATACCTCATATGGAAGGAAATCCTGCCCCACAAAG 5189
834 luGlyLeuGluThrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
5190 AAGGGTTAGAAATGTCCAACAGATACATCGTCTCACCCACCTAGAACT 5239
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
5240 AAACACTCGACAGTGTGTCAGAACATCCCTTATCATGTTCTGAGGCT 5289
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
5290 ACCAGAGTGGCTGACTCGGTGTCAAACATTTGTGTCCTGCCAGCTGG 5339
884 aAlaAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
5340 TTAATGCTAATCTTCCAGAAATACCTCCAGAAAGAGACTAAGGGGAAGC 5389
901 HisProGlyAlaHisTrpGluValAlaAspPheThrGluValLysProAlaLy 917
5390 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCGGCTAA 5439
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
5440 ATACGGAAACAATATCTATTGTTTGTGACACACCTTTTCAGGATGG 5489
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
5490 TAGAGCTTATCTACTAGAAAGAGACTTCAACCGTGTGCTAATAAAA 5539
951 ileLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
5540 ATACTGGAGAAATTTTCCGAGATTGGAATACCTAACCTAAGGTAATCGGTC 5589
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5590 AGCAATGGTCCAGCTTCTGTCGCCAGGTAAGTCAGGACGTGGCCAAGA 5639
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5640 TATTGGGATTGATTGAAACTGCATTGTGCATACAGACCCCAAGCTCA 5689
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5690 GGACAGGTAGAGAGTGAATAGAACCATTAAGAGACCCCTTACCAAAATT 5739
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
|||||

```

```

5740 GACCACAGAGACTGGCATTAATGATTGGATGGCTCTCTCCTCCCTTTGTGC 5789
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5790 TTTTTAGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5839
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
5840 TTGCTCTACGGGGGACCCCGCTTGGTAGAAATTCCTCTACATAG 5889
1067 rAlaAspValIleLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5890 TGCTGATGTGCTGCTTTCCAGCCTTTGTCTCTAGCTCAAGCGCTCG 5939
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5940 AGTGGGTGAGGCAACGAGCGTGGAAAGCAGCTCCGGGAGGCTTACTCAGGA 5989
1101 ...GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
5990 GAAGGAGACTTGCAAGTTCACATCGCTTCCAAGTTGGAGATTCACTTA 6039
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
6040 TGTTAGACGCCACCGTCGAGAAACCTCGAGACTCGGTGGAAGGCCCTT 6089
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
6090 ATCTCGTACTTTTGACCACACCAACGCGTGTGAAAGTCGAAAGAATCCCC 6139
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerG 1166
6140 ACCTGGATTCATGCATCCACGTTAAGCGCGCCACCTCCCGATTCCGG 6189
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
6190 GTGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 6239
1183 euValProTyrSerAsnAsnSer 1191
6240 TGGTTCTTACTCTGCAATAACTCC 6265

```

seq\_name: gb\_vi:PEN133818

```

seq_documentation_block:
LOCUS PEN133818 8750 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env
genes and LTR (class B, clone 43).
ACCESSION AJ133818
VERSION AJ133818.1 GI:6688949
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol
protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses; Retrovirdae; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 8750)
AUTHORS Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)
MEDLINE 20219394
REFERENCE 2 (bases 1 to 8750)
AUTHORS Toenjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,
Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225
Langen, GERMANY
FEATURES
source Location/Qualifiers
1..8750
/organism="porcine endogenous retrovirus"
/proviral
/strain="type C"

```

```
/db_xref="taxon:61673"
/clone="43"
1. .629
467
misc_feature
primer_bind
632. .649
/note="cap site"
/note="gly4"
1064. .2635
/gene="gag"
1064. .2635
/gene="gag"
/function="retroviral capsid"
/product="Gag protein"
join(1064. .2635,2639. .6223,6099. .8072)
/note="viral polyprotein"
/codon_start=1
/protein_id="CAB65341.1"
/db_xref="GI:6688950"
/translation="MGQVTPPLSLTDHWTVEVRSRAHNSLVQVKKGPWQTFCASEWP
TFDQWPSSEGTNSEIILAVKAIIFQTGPSSHPDQEPYILWQDLAEDPPWPWKPLN
KPRKPGPRILALGCKNKSAAKEVSPRIYPIETEPPTWPEQVPVPPYPAQGAVRG
PSAPGAPVVEGPAAGTSSRRGATPERTDEIAILPLRTYGPMPGQLOLPQYWPFS
ADLYNWKTNHPPSEDPORLTGLVESLFPESHQPTWDDCQQLLTFTTEERERILIEA
RKNVPGADRPQLQNEIDMGFPLTPGWDYNTAEGRESKIYRQALVAGUGASRRP
TNLAKVREVMQPNPEPSPVFLERLEAFRRPTPDPSTSEAKQASVALAFIQSALDIR
KKLQRLGLEAEALDLVREAEKVYRRTEEEKQREKEEREERERRDRROEKNLIT
KILAAVEGCKSRERERDFRKIRSGPROSGNLGNRTPLDKQOCAYCKKKGHWRNCPK
KGNKGPVLALEDKDRGRSDPLPEPRVALKVGQPVFELVDTCAEHSVLLOPLGKL
KEKSKVMGATGATGQRYPWMTTRTVDLAGRVTHSELVPECPVPLGRDILTKMGQI
SFEQGRPEVSVNNKPIVTLTQLDDEYRLYSPQVAPDQDIQSWLEQFPQAAETAGMG
LAKQVPPQIKASATPVSRQYPLSREAREGIWPHVORLTQOGLVYVQSPNNTP
LVPRKPNNDYRPVQDLREVNKRQVDIHTVFNPNYLLSALPPERNMYTVLDLDAF
CURLHPTSOPLFAFEWRDPTGRTGQLTWRLPOGFKNSPITFDEALHRLDLANRQIH
POVTLLOVDDLLLAGATKQDCLEGTKALLLESLDLYRASAKKQICRREVTYLGYS
LGGORWLEAEKTVQVLPATTAQVREFLGTAGFCRLWIPGEPATLAAPLYLPIKE
KGEFSAPEHQTPDAIKKALLSAPALPDVTKFTFLYVDERKGVARGVLTQTLGPW
RPAPVLSKKDQPVASGMPVCLKAIAVAAILVKDADKLTGCONITVIAPHALENIVRQ
PPDRWMTNARTHYQSLLTERTVFAAPALNPATLLPEETDEPTVHCHQLLIEETG
VRKDLTDLPTGEVLTWFTDGSYYVEGKRWAGAAVVDGTHTWASSLPEGTSAQAE
LMALTOALRLABKGSINIVTDSRYAFATAHVGALYKORGLLTSAGREIKNKEETLSL
LEALHLPRKLAIIHCPCGHQKADLISRGNOADRVAQAAQVNLPIIETPKAPEPR
RQYTLDEQWEIKIQDSETEPGTCITSTGKEILPHKSGLEYVOOIHRLTHLGRHLQ
QVTRSPHYULRPGVADSVKVCVPCQVKNANPSRVPVPPKRLGSHPGAHWEVDFT
PAPKARYKULLLVYDFVTSWVWEAYPTKKTSTVVAKKILEEIPFPIGKIVGSDNG
PAFVAQVSGKLAKILGIDWKLVCAVPOSSGQVEMNKTIKETILTKLTPTGINDWIA
LLPFVLRVNTPGQGLTPVLELLYGGPPPLVEIASVHADVLLSQPLFSLKALEWV
RORANKQLREAYSGEGDLQVPHRFQVGSVYVRHRAGNLETRKNGPYLVLLTPTAV
KVEGLSTWISHAVLAPPPDSGWRAEKTEPNLKLRLHRLVPIYNNNSPQGMHPTLSW
RHLPTRGEPKRLRILPLSFATAMFLTITPQASSKRLIDSNNPHRPLSLTLIIDP
DTGVTVNSTRGVAPRGTWPELHFCRLINPAVKSTPPNLVRSYGFYCCPGTEKEKY
GGSESCFRNWCSTNSDQWKPILSLODRVKEFVNSGPGKYKVMKLYKOKSCSPD
LDYLSIFTEKGQENIOKINGMSWGIYFYKGGAGSTLTLRLITGTETPPVAVG
PDKVLAEGPPALPEPHNLVPOLITSLRPDITQPPSNGTGLTIPNTPRNSPGVPVK
QORLPSLQGAFAALNSDTPDASCWLCSSGPPYEGMAKEGKFNVYKHEHNOCTW
GSRNKLTLTEVSGKTCGICAPPSSHLCYSTVVEQASENQYLPVGNRWACNTGL
TPCVSSVFNQSKFCVMQIVPVRYHYPEEVLDEVDYRYNRPKREPVSUTLAVMLG
LCTAVGCTGTAAALITPOOLEKGELHAAMTEDLRALESVSNLESLTSLSEVL
QNRGLGLDLFLREGGLCAALKECCFYVDHSCALRDSMSKLRELRERRRRREADQW
FKGWNRSPWMTLLSALTGLVLVLLLTVGPCLINRFVAFVRERVSVAOIMVLRQO
YQGLSQGETDL"
2639. .6223
/gene="pol"
/function="retroviral protease and polymerase"
/product="Pol protein"
2639. .6223
/gene="pol"
6099. .8069
/mat_peptide
/gene="env"
/function="retroviral envelope"
/product="Env protein"
6099. .8069
/gene="env"
8122. .8750
LTR
```

polyA\_signal 8643. .8648  
BASE COUNT 2288 a 2214 c 2286 g 1962 t  
ORIGIN

alignment\_scores:  
Quality: 6223.50 Length: 1195  
Ratio: 5.230 Gaps: 1  
Percent Similarity: 99.582 Percent Identity: 98.828

alignment\_block:  
US-09-171-553B-5 x PEN133818 ..  
Align seg 1/1 to: PEN133818 from: 1 to: 8750

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
2639 GGGAGACGGGGTTCGACCCCTCCCGAGCCCGAGGTAGCTTTGAAGGT 2688  
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
2689 GGAGGGGCAACCACTTGGAGTTCTGCTTATACCGGAGCGGAGCATTCAG 2738  
34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
2739 TGCTGCTACCACTTAGGAAACTAAAGAAAAAATCTTGGGTGATG 2788  
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAs 67  
2789 GGTGCCACAGGGCAACGGGAGTATCCATGGACTACCCGAAGAACCGTTGA 2838  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84  
2839 CTGGCAGCTGGGACGGGTAAACCCACTCGTTCTGGTCATCCCTGAGTGCC 2888  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
2889 CAGTACCCCTTCTAGTACAGACTTACTGACCAAGATGGAGCTCAAAAT 2938  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
2939 TCTTTTGAACAAGGAGACCAAGAGTCTGTGTAACAAACCAACCATCAC 2988  
117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134  
2989 TGTGTTGACCCCTCAATTAAGATGATGAATATCGACTATATTCTCCCAAG 3038  
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
3039 TAAAGCCTCATCAAGATATACAGTCTCTGTTGAGCAGTTTCCCAAGCC 3088  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
3089 TGGGCAAAACCGCAGGGATGGTTTGGCAAGCAAGTTCCTCCCAAGCT 3138  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
3139 TATTCACTGAAGGCCAGTGCCTACACAGTATCAGTCAGACAGTACCCT 3188  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeulle 200  
3189 TCAGTAGAGAGGCTCGAGAGGAATTTGGCCGCATGTTCAAGATTAATC 3238  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
3239 CAACAGGGCATCCTCTGCTGCTCAATCCCTTGGAACTACTCCCTGCT 3288  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
3289 ACCGGTTAGGAAGCTGGGACCAATGATTATCGACCACTACAGGACTTGA 3338  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250

3339 GAGAGTCAATAAAAGGGTGCAGGACATACACCCAAACGGTCCCGAACCCCT 3388  
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLe 267  
3389 TATAACCTCTTCCAGCGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 3438  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
3439 GGACTTAAAGATGCTTCTTCTGCTGAGATTACACCCACTAGCCAAAC 3488  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
3489 CGCTTTTTCCTTCCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3538  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
3539 CTCACCTGAGCCGACTGCCCCAAGGGTTCAAGAACTCCCGACCATCTT 3588  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3589 TGACGAAGCCCTACACAGGACCTGCCCAACTTCAGGATCCAAACACCCCTC 3638  
334 InValThrLeuLeuGlnTrpValAspLeuLeuAlaGlyAlaThr 350  
3639 AGGTGACCCCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3688  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuSerAs 367  
3689 AACAGGACTGCTTGAAGGTACGAAGGCATCTACTGCTGGAAATGTCCTGA 3738  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3739 CCTAGGCTACAGACCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAGG 3788  
384 aThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
3789 TAACATACCTGGGGTACAGTTTGCGGGGCGGCGCAGCGATGGCTGACGGAG 3838  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
3839 GCACGGAAGAAACTGTAGTCCAGATACCGGGCCCCCAACACACAGCCAAACA 3888  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
3889 AGTGAGAGAGTATTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 3938  
434 LpPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
3939 GGTTCGGACCTTAGCAGCCCCCACTCTACCCGCTAACCAAGAAAAGGG 3988  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
3989 GAATTTCTCTGGCTCCTCAGCACCAGAGACATTTGATGCTATCAAAA 4038  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProaspValThrLysProp 484  
4039 GGCCCTTGCTGAGCGACCTGCTTGGCCCTCCCTGACSTAACTAAACCCCT 4088  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
4089 TTACCCCTTATGTGGATGAGCGTAAAGGAGTAGCCCGGAGGATTTTAACC 4138  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
4139 CAAACCCCTAGGACCATGGAGGACCTGTTGCTACCTGCTCAAGAAGCT 4188  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534  
4189 TGATCTGTAGCCAGTGGTTGGCCCGGTATGCTGAAGGCTATCGCAGCTG 4238  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
4239 TGGCCATACTGGTCAGGACGCTGACAAATTCGCTTGGGACAGATATA 4288

551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
4289 ACTGTATAGCCCCCATGATTTGGAGAACATCGTTGGCAGCCCCCAGA 4338  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
4339 CCGATGATGACCAACGCCCGCATGACCCACTATCAAAAGCTGCTTCTCA 4388  
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600  
4389 CAGAGAGGTCACTTTCGCTCCACCGCGCTCTCAACCCCTGCCACTTT 4438  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
4439 CTGCTCAAGAGACTGATGAACCCAGTCACTCATGATTGCCATCACTATT 4488  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4489 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4538  
634 LpGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
4539 GAGAAGTGTCTAACTGCTTCTCAGCGAAGCAGCATGTTGGTGGAGGT 4588  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4589 AAGAGGATGCTGGGGCGCGGTGGTGCACGGGACCCACACGATCTGGGC 4638  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4639 CAGCAGGCTGCCGGAAGAACTTTCAGCGAAAAGGCTGAGCTCATGGGCC 4688  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4689 TCACCGAAGCTTTCGGCTGGCCGAGGAGAAATCCATAAATTTATACG 4738  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
4739 GACAGCAGTATGCTTTCGGACTGCACAGTACACGGGCGCATCTATAA 4788  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4789 ACAAGGGGGTGTCTTACCTCAGCGAGGAGGAAATTAAGAACAAAGAGG 4838  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4839 AAATTCAGGCTATTAGAACCTTACATTTGCCAAAAGGCTAGCTATT 4888  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4889 ATACACTGCTCTGGACATCAGAAAGCCCAAGATCTCATATCTAGAGGAA 4938  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
4939 CCAGATGGCTGACCGGGTGGCAACAGCAGGAGCCCGAGCTGTTAACCTTC 4988  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
4989 TGCCTATAATAGAAACGCCCAAGCCCAAGCCAGAACCCAGACAGTACACC 5038  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
5039 CTAGAAGACTGGCAAGAGATAAAAAGATAGACCAGTTCTCTGAGACTCC 5088  
817 oGluGlyThrCystThrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
5089 GGAGGGACCTGCTATACCTCATATGGAGGAAATCTGCCCCCACAAG 5138  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
5139 AAGGTTAGAAATATGTCACAGATACATGCTCTCACCCACCTAGGAAC 5188

```
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
5189 AACACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTCTTCAGGCT 5238
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
5239 ACCAGGAGTGGCTGACTCGGTGGTCAACAACTTGTGTCCTGCCAGCTGG 5288
884 aAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900
|||||
5289 TTAATGCTAATCTTCAGAGTACCTCCAGAAAGAGACTAAGGGGAGC 5338
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
5339 CACCCAGCGCTCACTGGGAGTGGACTTCACCTGAGGTAAAGCGGCTAA 5388
917 sTyrGlyAsnLysTrpLeuValPheValAspThrPheSerGlyTrpV 934
|||||
5389 ATACGGAACAATATCTATTGGTTTGTAGACACCTTTTCAGGATGGG 5438
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
5439 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 5488
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967
5489 ATACTGGAGAAAATTTTCCAGATTGGGAATACCTAAGTAATCGGGTC 5538
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5539 AGACAATGGTCCAGCTTTTTCGCCAGGTAAAGTCAGGACTGGCCAAGA 5588
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5589 TATTGGGATGATTGGAACACTGCATGTGCATACAGACCCCAAGACTCA 5638
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
5639 GGACAGGTAGAGGATGAATAAACCATTAAGAGAGACCTTACCAAAAT 5688
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034
5689 GACCAGAGACTGGCATTAAATGATTGATAGTCTCTCCCTGCGCTTTGTC 5738
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5739 TTTTATAGGTAGACACACCCCTGGACAGTTGGCGCTGACCCCTATGAA 5788
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
5789 TTGCTCTACGGGGGACCCCGCTGGTAGAAATTTGCTTCTGATACATAG 5838
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5839 TGTGTATGTCGCTTTCACGCTCTGTTCTCTAGCTCAAGGGGCTCG 5888
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
5889 AGTGGGTAGGCACACAGCGTGAAGCAGCTCCGGAGGCGCTACTCAGGA 5938
1101 ...GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
5939 GAAGGAGACTTGCAGCTTCCACATCGCTTCCAAGTGGGAGATTCAGTCTA 5988
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProF 1133
5989 TGTATAGCCCGCTGCGAGAAACCTCGAGACTCGGTGGAAGGCCCTT 6038
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
6039 ATCTCGTACTTTTGACCACACAGCGCTGTGAAAGTGAAGGAATCTCC 6088
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG 1166
```

```
|||||
6089 ACCTGCATCATCATCCACCGTTAAGCTGGCGCCACCTCCCGACTCGG 6138
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
6139 GTGGAGAGCGGAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCC 6188
1183 euValProTyrSerAsnAsnSerProGlyGln 1194
6189 TGGTTCCTTACTCTACAATAACTCCCGAGGCCAG 6223
```

seq\_name: gb\_v1:PERY17013

seq\_documentation\_block:

LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000  
DEFINITION Porcine endogenous retrovirus gag, pol, env genes.

ACCESSION Y17013

VERSION Y17013.1 GI:3005985

KEYWORDS env gene; gag gene; pol gene.

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

Viruses; Retroviral viruses; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 7808)

AUTHORS Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.

TITLE Establishment and characterization of molecular clones of porcine

endogenous retroviruses replicating on human cells

J. Virol. 74 (9), 4028-4038 (2000)

MEDLINE 20219394

REFERENCE 2 (bases 1 to 7808)

AUTHORS Toenjes,R.R.

TITLE Direct Submission

JOURNAL Submitted (30-MAR-1998) R.R. Toenjes, Paul-Ehrlich Institut,

Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG

LOCATION/Qualifiers

1..7808

/organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"

/tissue\_type="kidney"

/cell\_line="PK15"

/gene="gag"

/gene="gag"

/gene="gag"

/codon\_start=1

/protein\_id="CAA76581.1"

/db\_xref="GI:3005986"

/db\_xref="SPTREMBL:O73504"

/translation="MSQTVTTPLSLTLDHWTVEVRSRAHNSVQVKKGPMQTCASEWP

TQVWPSSEGTNSEIILAVKALIFQTPSSHPDQEPYLLTWQDLAEDPPVWKPLN

KPKPGPRILALGKKNKHAKEVPSRIPEIEEPTWPEQPPVPPYPAQAVRG

PSAPPVVEGPAAGTSGRGATPERTDEIALLRTYVPPMPGQLOPLQWPFSS

ADLYNMKTNHPPTSEDQRLTGLVESLMFSGQPTWDCQQLQTLTFAERERILLEA

RKNVPGADRGPTQLQNEIDMGFPLTRLDWDYNTAEGRESLKIYRQALVAGLRGASRRP

TNLAKREVQGNPEPVSVELERLMEAFRRFTPEPTSEAKASVALAFIGOSALDIR

KKLORLEGLQEAFLRDLVBEAKYVGRETESEKEREEREERDRQEKNLIT

KILAAVVEKSSRRERDRFRKIRSGPRQSGNLNRTPLDKDQCAICYCKEGRHARNCPK

KNGKGPVLALEEDKD"

2120..4726

/gene="pol"

/gene="pol"

/codon\_start=1

/protein\_id="CAA76582.1"

/db\_xref="GI:3005987"

/db\_xref="SPTREMBL:O73505"

/translation="MCAQISFEQGRPEVSVNNKPTITLTQLDDEVRILYSPQVKPDQD

IQSWLEQFPQAAETAGMGLAKQVPPVITQLKASATPVSVROYPLSRARAGINPHVQ

RLIQGILYVQSPWNTPLLPVRKGTNDIRVQDLREKNRKVDIHTVTPNPNLLS

ALPPERNWYTVILDKDAFCLRLHPTSQPLFAFEWRDPTGQTGTLTTLPLPQGFNS

PTIFDEALHRDLANFRIOHPQVTLQYVDDLLLAGATQDCLEGTAKLLILSLDKYR

ASAKKAQICRREVTVLIGSRGQRLTEARKKTVVQIPAPTTAKQVREFTGTAGFCR

LWIPGFATLAAPLYPLTKREKGFSSNAPEHQAFDAIKALLSAPALALPDVTKPTLY

```

VDRKRGVARGVLTQTIGPWRPVPVAVLSKLLDPVAGSWPVCLKAI AAVAILVKDADKLT
LGNITVIVADHENIVRQPPDRWMTNARMTHYQSLLLTERTVFAPPAALNPAFLTLE
ETDPEYHAPHCHOLLIEETGRVKDLTDIPLTGEVLTWFTDSSYVVEGGRMAGAAVVDG
THIWIASSPEGTSAQAEALMALTOALRLAEGKSINIYDTSYAFATAHVHCAIVKOR
GLTISAGREIKNEEELSLEELHLPKRLAIITHCPOHAKDLISRGNOQMDRVAKQA
AQAVNLLPIETPRAPRPRQYTTLEDQWEEKIDQFSETPGETSYSGKEILLPHKEG
LEYVOQIHLRTHLTKLQOLVRYTSPYHVRLEPVGADSVYKHCPCOLVNPNSRIPP
GKRLRSGHPGAHWEVDETEVKPAKYNKYLLVFEVDTFSGWEAYPTKKTSTVVAKKY
WRKFFRDLVLR"
5294..7267
/genes="env"
CDS
5294..7267
/genes="env"
/codon_start=1
/protein_id="CAA76583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:O73506"
/translation="MHPTLSWRHLPTGGEGPKRLRIPLSFASIANPLILITITPQASSK
RLDSSNHRPLSLTWLIIDPDGTVTNSTRGVAGRTWMPHELFCRLINPAVKSTP
PNLVSRYGFCVCPGTEKRYKESGSEFCRWSCVTSNDGDKWPISLQDRVKFSEVN
SGPKYKVMKLYKDKSCPSDLIDLYKISFTEKGKQENIOKWTNGMSWGIYFYKYGGGA
GSLTLRLLETGTETPPVAVGPDKVLAEQGPALPPHNLVPOLTSRLPDTTOPPSN
GTLTLPTNTPNSPCVPVKTGQRLFSLIQARFQALNSTDPDATSCWLCILSSGPPY
EGMAKSGKNVTKEHRNCTWGSRNKLTITLTVSGKGTGIGKAPPSSHQHLCTSYVIEQ
ASENQYLVPYNRWACNTGLTFCVSSSVFNQSKDFVMVQIVPRVYHPEEVILDEY
DYRNRKREPVSLLTSLVMLGLTAVGVTGTAALITGPQOLEKGLGELHAAAMTDLR
ALBESVNLPELSTLSVVLQNRGLDLFLREGGLCAALKKEECFCFYVDHSGAIRDS
MSKLRELRERREREADOGWFEWPNRSPWMTLLSALTGLPLVLLLLLTGVPCLIN
RFVAFVREVRGAVQIMVLRQOYQGLLSQGETDL"
BASE COUNT 2077 a 1971 c 2047 g 1713 t
ORIGIN

```

```

alignment_scores:
  Quality: 6202.50      Length: 1195
  Ratio: 5.217         Gaps: 1
  Percent Similarity: 99.498  Percent Identity: 98.912

alignment_block:
US-09-171-553B-5 x PERV17013 ..
Align seg 1/1 to: PERV17013 from: 1 to: 7808

1  GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
1835 GGGAGACGGGGTTCGGACCCCTCCCTGAGCCAGGGTAACTTTGAAGGT 1884

17  lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34
|||||
1885 GGAGGGGCAACAGTGTAGTTCCTGGTATACCGGAGCGGAGCAATTCAG 1934

34  alLeuLeuGlnProLeuGlyLysLysGluLysLysSerTrpValMet 50
|||||
1935 TGCTGCTACACCATTAAGAAAACCTAAAGAAAAAATCCTGGGTGGTG 1984

51  GlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgThrValAs 67
|||||
1985 GGTGCCACAGGCAACGGCAGTATCCATGGACTACCCGAAGAACCGTTGA 2034

67  pLeuGlyValGlyArgValThrHisSerPheLeuValTlleProGluCysp 84
|||||
2035 CTTGGCAGTGGGACGGGTAAACCATCGCTTCTGCTCATCCCTGAGTGCC 2084

84  roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnTlle 100
|||||
2085 CAGTACCCCTCTTAGGTAGAGACTTACTGACCAAGATGGAGCTCAAT 2134

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2135 TCTTTTGAACAAGGAAGACAGAAAGTGTCTGTGAATAACAACCCATCAC 2184

117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
|||||

```

```

2185 TGTGTTGACCCCTCCAATTAGATGATGAATATACGACTATATTTCTCCCAAG 2234
134 allLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2235 TAAAGCCTGATCAAGATATACAGTCTCTGGTTGGAGCAGTTTCCCAAGCC 2284
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2285 TGGGCAGAAACCCAGCGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 2334
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2335 TATTCAACTGAAGCCAGTGTACACCATGATCAGTCAGACAGTACCCCT 2384
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuL 200
|||||
2385 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAAATC 2434
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuL 217
|||||
2435 CAACAGGGCATCTAGTTCCTGTCCTCAATCCCTTGGAACTATCCCTGCT 2484
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2485 ACCGTTTAGAAGCCCTGGGACCAATGATATATCACCAGTACAGGACTTGA 2534
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2535 GAGAGGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCCCT 2584
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
|||||
2585 TATAACCTCTTGAGCGCCCTCCCGCTGAAACGGAACCTGGTACACAGTATT 2634
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2635 GGACTTAAAGATGCTCTCTCTGCTGAGATTACACCCACTAGCCAAC 2684
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
2685 CACTTTTGGCTTTCGAATGGAGATCCAGGTACGGGAAGAACCGGGCAG 2734
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
2735 CTCACCTGGACCCGACTGCCCAAGGTTCAAGAACTCCCCGACCATCTT 2784
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
2785 TGAGAAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAACACCCCTC 2834
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
2835 AGGTGACCCCTCTCCAGTAGTGGATGACCTGCTTCTGGCGGGAGCCACC 2884
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAs 367
|||||
2885 AAACAGACTGCTTAGAAGGTACGAGGCACTACTGCTGGAATTTGCTGA 2934
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
2935 CCTAGGCTACAGACCTCTGCTAAGAGGCCAGATTTCAGGAGAGAGG 2984
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
2985 TAACATACTTGGGGTACAGTTTGGGGGGGGGCGGAGCGATGGCTGACGAG 3034
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
3035 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACAGCCARACA 3084
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3085 AGTCAGAGAGTTTTGGGGGACAGCTGGATTTTGCAGACTCTGGATCCCGG 3134

```

434 lypheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
3135 GGTTCGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGGG 3184  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
3185 GAATTCCTCTGGGCTCTGAGCACCAGAGGCGATTTTGATGCTATCAAAA 3234  
467 sAlaLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484  
|||||  
3235 GGCCTCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAACCCCT 3284  
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500  
|||||  
3285 TTACCCCTTTATGTGATGAGCGTAAGGAGTAGCCCGAGGAGTTTAAACC 3334  
501 GlnThrLeuGlyProTrpArgProValCysLeuLysAlaIleLysLysLe 517  
|||||  
3335 CAACCCCTAGCACCATGGAGGAGACCTGTTCCTACCTGTCAAGAGAGCT 3384  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlay 534  
|||||  
3385 TGATCTCTAGCCAGTGTGGCCGATGCTGAAGGCTATCGCAGCTG 3434  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
3435 TGGCCATACCTGGTCAAGSACGCTGACAAATTGACCTTGGGACAGATATA 3484  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
3485 ACTGTAATAGCCCCCATGATGATGGAGAACATGTTTCGGCAGCCCCCAGA 3534  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
3535 CCGATGATGACCAACCCCGCATGACCCACTATCAAGGCTGCTCTCA 3584  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
3585 CAGAGAGGGTCACCTTCCTGCCACCAGCCGCTCTCAACCCCTGCCACTCT 3634  
601 LeuProGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
3635 CTGCCTGAAGAGACTGATGAACACGAGTGCATCATTCATGCCATCAACTAT 3684  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
3685 GATTGAGGAGACTGGGTCTCCGCAAGGACCTTACAGACATACCCGCTGACTG 3734  
634 yGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
3735 GAGAAGTCTTAACCTGTCTCACTCAGCGAAGAGCTATGTGGTGAAGGT 3784  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
3785 AAGAGGATGCTGGGGCGGGGTGGTGGACGGGACCCACACGATCTGGGC 3834  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
3835 CAGCAGCCTGCCGAAGGAACCTTACGGCGCAAAAGGCTGAGCTCATGGCCC 3884  
684 euthrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
3885 TCAGCGAAGCTTTCGGCTGCGCGAAGGAAATCCATAACATTTATACG 3934  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
3935 GACAGCAGGTATGCTTTTGGGACTGTCACACGTACACGGGGCCATCTATAA 3984  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
3985 ACAAGGGGGTGTCTTACCTCAGCAGGGGGGAAATAAGAACAAAGAGG 4034

734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4035 AAATTCCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGTATT 4084  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
4085 ATACACTGTCTTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4134  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
|||||  
4135 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGCTGTAACTTTC 4184  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
4185 TGCCTTATAAGAAAGCCAGAGCCCAAGAACCCAGACAGACAGTACACC 4234  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
4235 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGACTCC 4284  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
4285 GAGGGGACCTGCTATACCTCATATGGGAAGGAATCTGCCCCACAAG 4334  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
4335 AAGGGTTAGAAATATGTCCAACAGATACATCGTCTCACCCACCTAGAACT 4384  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
4385 AAACACCTGCAGCAGTTGCTCAGAACATCCCTTATCATGTTCTGAGGCT 4434  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
4435 ACCAGAGTGGCTGACTCGGTGTCAACATTTGTGTGCTGCCAGCTGG 4484  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
4485 TTAATGCTAATCTTCCCAAAATACCTCCAGGAAGAGACTAAGGGAAGC 4534  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
4535 CACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 4584  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
4585 ATACGGAACAAATATCTATTGTTTTGTAGACACCTTTTCAGGATGG 4634  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
4635 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 4684  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
4685 TA CTGGAGGAAATTTTCCGAGATTTGGAAATACCTAAGGTAAATCGGGTC 4733  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
4734 ACACAAATGGTCCAGCTTTTGTGGCCAGGTAAGTACAGGACTGGCCAAGA 4783  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
4784 TATTGGGATTTGATTGGAACCTGCATTTGTGCATACAGACCCCAAGCTCA 4833  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
4834 GGACAGGTAGAGAGGTGAATAGAACCATTAAAGAGACCTTTACCAAAAT 4883  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
4884 GACCACAGAGACTGGCATTAAATGATGGATAGCTCTCTCTGCTTTGTCG 4933  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050





LINRFVAFVRQVSQSAVRIMWLRLQQVQGLPS"  
LTR 8216..8918  
polyA\_signal 8810..8815  
BASE COUNT 2435 a 2174 c 2253 g 2056 t  
ORIGIN

## alignment\_scores:

Quality: 6142.00 Length: 1191  
Ratio: 5.196 Gaps: 0  
Percent Similarity: 99.244 Percent Identity: 97.565

## alignment\_block:

US-09-171-553B-5 x PEN293656 ..

Align seg 1/1 to: PEN293656 from: 1 to: 8918

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuIysVa 17  
|||||  
2728 GGAAGACGGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACCTTTAAAGGT 2777  
17 lGluGlyInProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
|||||  
2778 GGAGGGGCACACCAAGTTGAGTTCTGTTGATACCGGAGCGAAACATTCAG 2827  
34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysLysSerTrpValMet 50  
|||||  
2828 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2877  
51 GlyAlaThrGlyGlnArgGlnTrpProTrpThrArgArgThrValAs 67  
|||||  
2878 GGTGCCACAGGGCAACCAATAATCCATCCGACTACCGGAAGACAGTTGA 2927  
67 pLeuGlyValClyArgValThrHisSerPheLeuValIleProGluCysP 84  
|||||  
2928 CTTGGGAGTGGGAGGGTAACCCACTCGTTCTGGTCATACCTGAGTGCC 2977  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
|||||  
2978 CAGCACCCCTCTTAGGTAGAGACTTACTGACCAAGATGGGAGCACAAATT 3027  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
|||||  
3028 TCCTTTGAACAAGGGAAACACAGAGTCTCTGCAATAAACAAACCTATCAC 3077  
117 rValLeuThrLeuGlnLeuAspAspGluTyArgLeuTySerProGlnV 134  
|||||  
3078 TGTGTTGACCCCTCCCAATTAGATGACGAATATCGACTATATTCCTCCCTAG 3127  
134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
|||||  
3128 TAAAGCCTGATCAAAATATACAAATCTGTTGGAACAGTTTCCCAAGCC 3177  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
|||||  
3178 TGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 3227  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTrpProL 184  
|||||  
3228 TATTCACCTGAAGCCAGTGTCTGCACCAAGTGTCACTGACAGTACCCCT 3277  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
|||||  
3278 TGAGTAAAGAAGCTCGAAGGAATTCGGCGGCATGTTCAAAAGATTAAATC 3327  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
|||||  
3328 CAACAGGGCATCCTAGTTCCTGTCCTCAATCTCCCTGGAATACCTCCCTGCT 3377  
217 uproValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
3378 ACCGGTTAGAAAGCCCTGGGACTAATGACTATTCGACCACTACAGCAGCTGA 3427

234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
3428 GAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAACCT 3477  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
|||||  
3478 TATAACCTCTTGTGTCTCCACCCCAACGAGCTGTTATACAGTATT 3527  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
|||||  
3528 GGACTTAAAGGATGCTTCTTCTGCTGAGACTACATCCCACTAGCCAAAC 3577  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3578 CACTTTTTCCTTCGAATGGAGATCCAGGTGCGGGAAGAACCGGGCAG 3627  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3628 CTCACCTGGACTCGACTGCCCAAGGGTTCAAAAACCTCCCGACCATCTT 3677  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3678 TGACGAGGCCCTACACAGAGACTGGCCAACTTCAGGATCCCAACACCCC 3727  
334 lnValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThr 350  
|||||  
3728 AGGTGACCCCTCCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGGCCACC 3777  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
|||||  
3778 AAACAGGACTGCTTAGAGGTACGAGGACACTACTGCTGGAATTGTCTGA 3827  
367 pLeuGlyTyArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
|||||  
3828 CCTAGGCTACAGAGCTCCGCTAAGAAGGCCCAAGATTTGCAGAGAGAGG 3877  
384 alThrTyLeuGlyTySerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
|||||  
3878 TAACATACTTGGGTACAGTTTTGCGGGCGGCGAGCGATGGCTGACGGAG 3927  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417  
|||||  
3928 GCACGGAAGAGAACTGTAGTCCAGATACCGGCCCAACCCACAGCCAAACA 3977  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
3978 AGTGAGAGAGTTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 4027  
434 lypPheAlaThrLeuAlaAlaProLeuTyProLeuThrLysGlyLysGly 450  
|||||  
4028 GGTTCGACCTTAGCAGCCCTCTACCTACCACTAACCAAGAAAGAGG 4077  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
4078 GAATTCCTCTGGGCTCCTGAGCACCAAGAGGCAATTTGATGCTATCAAAA 4127  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
4128 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTGACTAAACCT 4177  
484 heThrLeuTyValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
4178 TTACCCCTTTATGTGGATAGCGTANGGAGTAGCCCGGGAGTTTTTAACC 4227  
501 GlnThrLeuGlyProTrpArgProValAlaTyLeuSerLysLysLe 517  
|||||  
4228 CAAACTCTAGACCATGGAGAGACCTGTTCCTTACCTGTCAAGAGACT 4277  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534  
|||||  
4278 CGATCTCTGTAGCCAGTGGTTGGCCCATATGCTTGAAGGCTATCGCAGCTG 4327  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550

|||||  
4328 TGCCATACCTGGTCAAGGACGCTGCACAAATTGACTTTGGACAGAAATATA 4377  
|||||  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
4378 ACTGTAATAGCCCCCATCGGTGGAGACATCGTTCCGACGCCCCAGA 4427  
|||||  
567 pArgTTPMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
4428 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTCTCA 4477  
|||||  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
4478 CAGAGAGGTGACGTTCCGCTCCACCGAGCGCTCTCAACCCCTGCCACTCTT 4527  
|||||  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
4528 CTGCCTGAAGAGACTCATCAACCACTGACTCATGATGTCATCACTATT 4577  
|||||  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4578 GATTGAGGAGACTGGGCTCGGAAGACCTTACAGACATACCGCTGACTG 4627  
|||||  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4628 GAGAAATGTTAACCTGGTTCACTGACGGAAGCAGCTATATGTGGAAGGT 4677  
|||||  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4678 AAGAGGATGGCTGGGCGCGGTGGTGGACGGGACCGCCGACGATCTGGGC 4727  
|||||  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
4728 CAGCAGCCTGCGGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGCCC 4777  
|||||  
684 euThrClnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4778 TCACGCAAGCTTTGGCGGTGGCCGAAGGGAAATCCATAAACATTTATACA 4827  
|||||  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4828 GACACAGGTATGCCTTGGCAGCTGCACACGTACACGGGGCCATCTATAA 4877  
|||||  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4878 GCAAAAGGGGGTGGCTTACTCTCAGCAGGGAGGGAAATAAAGAACAAAGAG 4927  
|||||  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4928 AAATTCTAAGCCTATTAGAACCTTTACATTTGCCAAAAGGCTAGCTATT 4977  
|||||  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
4978 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 5027  
|||||  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
|||||  
5028 CCAGATGGCTGACCGGGTGGCAACGAGCGAGCGCCAGGCTGTTAACCTTC 5077  
|||||  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
5078 TGCCCTATGATGAACACCCCAAGGCCCAAGACCCCGGACGACGATACACC 5127  
|||||  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5128 CTAGAAGACTGGCAAGAAATAAAAAAGATAGACCAGTTCTCTGAAACTCC 5177  
|||||  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
5178 GGAGGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 5227  
|||||  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||

5228 AAGGTTTAGAATATGTCCAACAGATACATCGTCTAAACCCACCTAGGAAC 5277  
|||||  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
5278 AAACACCTGGACGAGTGGTTCAGAACATCCCTTATCATGTTCTGAGGCT 5327  
|||||  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
5328 ACCAGGAGTGCTGACTCGGTGGTCAAACTTGTGTGCCCTGCCAGCTGG 5377  
|||||  
884 alaAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
5378 TTAATGCTTAATCCTTCCAGATACCTCCAGGAAAGAGACTAAGGGGAGC 5427  
|||||  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
5428 CACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5477  
|||||  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
5478 ATACGGAACAAATATCTATTGGTTTGTGTAGACACCTTTTCAGGATGGG 5527  
|||||  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
5528 TAGAGGCTTATCCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5577  
|||||  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
5578 ATACTGGAGGAAATTTTCCAAAGATTTGGAAATACCTAAGGTATAGGGTC 5627  
|||||  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5628 AGACAATGGTCCAGCTTTCGTGCCAGGTAAAGTCAGGGACTGCCCAAGA 5677  
|||||  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
5678 TATTGGGGATTTGATTGGAAATCGCATTTGTGCATACAGACCCCAAGCTCA 5727  
|||||  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
5728 GGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGACCCCTTACCAAT 5777  
|||||  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
5778 GACCACAGAGACTGGCATTAAATGATTGGATGGCTCTCCTGCCCTTTGTGC 5827  
|||||  
1034 euPheArgValArgAspThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
5828 TTTTGGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5877  
|||||  
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5878 TTGCTCTACGGGGACCCCGCTGGCAGAAATTTGCCTTTGCACATAG 5927  
|||||  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
5928 TGCTGATGTGCTGCTTCCAGCGCTTTGTCTCTAGGCTCAAGGCGCTCG 5977  
|||||  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
|||||  
5978 AGTGGGTGAGGCAGCGAGCTGGAAGCAGCTCCGGGAGGCGCTTACTCAGGA 6027  
|||||  
1101 GlyAspLeuGlnValProHisArgPheGlnValIleGlyAspSerValTyrVa 1117  
|||||  
6028 GGAGACTTGCAAGTTCCACATCGCTTCCAAAGTTGGAGATTCAGTCTATGT 6077  
|||||  
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134  
|||||  
6078 TAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAAGGACCTTATC 6127  
|||||  
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150  
|||||  
6128 TCGTACTTTTGACCAACACCAACGCGCTGTGAAGTCGAAGAAATCCCCACC 6177

1151 TrpIleHisAlaSerHisValLysProAlaProProProAspSerGlyTr 1167  
 1167 TGGATCCATGATCCACAGTTAAGCGCGCCACCTCCCGATTCGGGGTG 6227  
 1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184  
 6228 GAAAGCGAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCCGCGTGG 6277  
 1184 alProTyrSerAsnAsnSer 1191  
 6278 TTCCTTACTCTGTCATAACTCC 6300

seq\_name: gb\_pat:AR130475

seq\_documentation\_block: AR130475 8132 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 3 from patent US 6190861.

ACCESSION AR130475

VERSION AR130475.1 GI:14118800

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 8132)

AUTHORS Fishman,J.A.

TITLE Molecular sequences of swine retroviruses method of using

JOURNAL Patent: US 6190861-A 3 20-FEB-2001;

FEATURES Location/Qualifiers

source

1..8132

/organism="unknown"

BASE COUNT 2248 a 1977 c 2037 g 1870 t

ORIGIN

alignment\_scores:

Quality: 6099.50 Length: 1191

Ratio: 5.173 Gaps: 1

Percent Similarity: 98.992 Percent Identity: 96.893

alignment\_block:

US-09-171-553b-5 x AR130475 ..

Align seg 1/1 to: AR130475 from: 1 to: 8132

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

1167 TGGATCCATGATCCACAGTTAAGCGCGCCACCTCCCGATTCGGGGTG 6227

2160 GGGAGACGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACCTTTGAAGGT 2209

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34

2210 GGAGGGCAACCAAGTTGAGTTCCTGTTGATACCGGAGCGAATTCAG 2259

34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50

2260 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2309

51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgThrValAs 67

2310 GGTGCCACAGGGCAACACAGTAGTATCCATGGACTACCGGAAGACGTTGA 2359

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84

2360 CTTGGGAGTGGGAGGGGTAAACCACTCGTTCTTGGTCATACCTAGTGCC 2409

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100

2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2459

101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleth 117

2460 TCTTTTGAACAGGGGAACAGAGAGTGTCTGCAAAATAAACACCTATCAC 2509

117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134

2510 TGTGTTGACCTCCCAATTAGATGACGAATATCGACTATACTCTCCCTAG 2559

134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150

2560 TAAAGCTGATCAAAATATACAAATCTGTTGGAACAGTTTCCCAAGCC 2609

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167

2610 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAGT 2659

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

2660 TATTCAACTGAAGCCAGTGCACACCACTGTCAGTCAGACAGTACCCCT 2709

184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200

2710 TCAGTAAAGAAGCTCAAGAAGGAATTCGGCCGCATGTCCAAAGATTAATC 2759

201 GlnGlnGlyIleLeuValProValGlnSerProTyrAsnThrProLeuLe 217

2760 CAACAGGGCATCTAGTCTCTGTCATCTCCCTGGATATCTCCCTGCT 2809

217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234

2810 ACCGGTTAGAAAGCCTGGGACTTAATGACTATCGACAGTACAGGACTTGA 2859

234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250

2860 GAGAGGTCAATAAAGGGTGCAGGATATACACCAACAGTCCCGAACCT 2909

251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValIle 267

2910 TATAACCTCTTGTGCTCTCCCAACCGAGCTGGTATACAGTATT 2959

267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284

2960 GGACITAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCAAAC 3009

284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300

3010 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3059

301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317

3060 CTCACCTGGACCCGACTGCCCAAGGTTCAAGAACTCCCGGACCATCTT 3109

317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334

3110 TGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCCAACACCTC 3159

334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350

3160 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGGCCACC 3209

351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367

3210 AAACAGGACTGTAGAGGACGAGGACACTTACTGCTGGAATTTGCTGGA 3259

367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384

3260 CTAGGCTACAGAGCTCTGCTTAAAGAGCCCGAGATTTGCAGAGAGAGG 3309

384 aIThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400

3310 TAACATACTTGGGTACAGTTTGGCGGACGGCGAGCGATGGCTGACGGAG 3359

401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417

3360 GCACGGAGAAAACCTGATGTCAGATACCGGCCCAACACACAGCAACA 3409

417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434

|||||  
3410 AATGAGAGAGTTTGGGACAGCTGGATTTTGCAGAGCTGGTGGATCCCGG 3459  
434 lypPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLysGly 450  
3460 GGTTCGACACCTTACGACCCCACTCTACCCGCTAACCAAGAAAAGGG 3509  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
3510 GAATTCCTCTGGGCTCCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 3559  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
3560 GGCCCTGTGAGCGACAGCTGCTGGCCCTCCCTGACGTAACAAACCCCT 3609  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
3610 TTACCCCTTATGTGGATGAGCGTAGGGAGTAGCCCGGGAGTTTAAACC 3659  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
3660 CAAACCCCTAGGACCATGAGAGAGACCTGTCGCCCTACCTGTCAAAAGAGCT 3709  
517 uasProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLay 534  
3710 CGATCCTGTAGCCAGTGTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyClnAsnIle 550  
3760 TGGCCATACTGGTCAAGGACCTGACAAATTTGACTTTGGGACAGAAATA 3809  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567  
3810 ACTGTAATAGCCCCCATGCAATTTGGAGAACATCGTTGGCAGCCCCAGA 3859  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrClnSerLeuLeuLeu 584  
3860 CCGATGGATGACCAACGCCGCTGACCCCACTATCAAGCCCTGCTTCTCA 3909  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
3910 CAGAGAGGTCAGCTTCCTCCACGAGCGCTCTCAACCCCTGCCACTCTT 3959  
601 LeuProGluGluThrAspGluProValThrHisaspCysHisGlnLeuLe 617  
3960 TTGCTGAAGAGACTGATGAACCAAGTGAATCATGATTCATCAACTATT 4009  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4010 GATTGAGGAGACTGGGTCGCCAAGGACCTTACAGACATACCGCTGACTG 4059  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValIleGluGly 650  
4060 GAGAAAGTCTAACCTGGTTCACTGACGGAAGCAGCTATGTGGTGAAGGT 4109  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4110 AAGAGGATGGTGGGGCGCGGTGGTGGACGGGACCCCGACAGTCTGGGC 4159  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4160 CAGCAGCCTGCCGGAAGAACTTACAGCAAAAAGGCTGAGCTCATGGCCC 4209  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4210 TCACGCAAGCTTGGCGGTGGCCGGAAGGAAATCCATAAATTTATACG 4259  
701 aspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
4260 GACACAGGTATGCTTTGGCAGCTGCACACGTACATGGGGCCATCTATAA 4309  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||

4310 ACAAGGGGGTGTCTTACCTACGAGGGAGGAAATAAAGAAAGAGG 4359  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4360 AATTTCTAAGCCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTATT 4409  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4410 ATACACTGTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAA 4459  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
4460 CCAGATGGCTGACCGGTTGCCAAGCAGGAGCCCGAGGTTGTTAACTTC 4509  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
4510 TGCCCTATATAGAAATGCCAAAGCCCCAGAACCCAGACGACAGTACACC 4559  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
4560 CTAGAAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTCAGACTCC 4609  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
4610 GGAAGGACCTGTCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAA 4659  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
4660 AAGGTTAGAAATATGCCAACAGATACATCGTCTAACCCACCTAGGAAC 4709  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
4710 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 4759  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
4760 ACCAGAGTGGCTGACTCGGTGTCACAACTTGTGTGCCCTGCCAGCTGG 4809  
884 aAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
4810 TTAATGCTAATCCTCCAGAAATGCCCTCCAGGAAGAGACTAAGGGGAAG 4859  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
4860 CACCCAGGCGCTCCTGGGAAGTGGACTTCACGTAGGTAAGCCGCTAA 4909  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
4910 ATACGGAACAATACCTATTGCTTTTGTAGACACCTTTTCAGGATGG 4959  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
4960 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCCGTGGTGGCTAAAA 5009  
951 IleLeuGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
5010 ATACTGGAGAAATTTTCCAGATTGGAAATACCTTAAGGTAATAGGGTC 5059  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5060 AGACAATGGTCCAGCTTTGTTCCCCAGGTAAAGTACGGACTGGCCCAAG 5109  
984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5110 TATTTGGGATTTGATTGGAACCTCATTTGTGCATACAGACCCCAAGCTCA 5159  
1001 GlyGluValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
5160 GGACAGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTAAAT 5209  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034  
5210 GACCCGGAGACTGGCGTTAATGATTGGATAGCTCTCTGCCCTTTGTGC 5259

1034 euPheArgValArgAsnThrProGlyClnPheGlyLeuThrProTyrGlu 1050  
|||||  
5260 TTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5309  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5310 TTACTCTACGGGGACCCCCCATTTGGTAGAAATTCCTTGTGTACATAG 5359  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
5360 TGCTGACGTGCTGCTTCCAGCCTTTGTCTCTAGGCTCAAGGCACITG 5409  
1084 lnrPValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
|||||  
5410 AGTGGGTGAGACAAACGAGCGTGGAGGCAACTCCGGGAGGCCTACTCAGGA 5459  
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValty 1116  
|||||  
5460 GGAGGAGACTTCAGATCCACATCGTTTCCAAAGTGGGAGATTCAGTCTA 5509  
1116 rValArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133  
|||||  
5510 CGTTAGACGCCACCGTCCAGGAAACCTCGAGACTCGGTGGAAGGCCCTT 5559  
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
|||||  
5560 ATCTCTACTTTTGACCACCAACCGGCTGTGAAGTCCGAAGGAATCTCC 5609  
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166  
|||||  
5610 ACCTGGATCATGATCCACCGTTAAACCGGCGCCACCTCCCGATTCGGG 5659  
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
|||||  
5660 GTGGAAGCGGAAAGACACTGAAATCCCTTAAGCTTCGCCCTCCATCGCG 5709  
1183 euValProTyrSerAsnAsn 1190  
:::|||||  
5710 TGGTTCCTTACTCTGTCAATAAC 5732

seq\_name: gb\_pat:AX052635

seq\_documentation\_block:  
LOCUS AX052635 4402 bp DNA PAT 12-JAN-2001  
DEFINITION Sequence 20 from Patent WO0071726.  
ACCESSION AX052635  
VERSION AX052635.1 GI:12226825  
KEYWORDS  
SOURCE  
ORGANISM  
porcine endogenous retrovirus.  
Viruses; Retroid viruses; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.  
1 (bases 1 to 4402)  
Fedorpiel,M.J.  
Methods to inhibit infectious agent transmission during  
xenotransplantation  
Patent: WO 0071726-A 20 30-NOV-2000;  
Mayo Medical Ventures (US)  
FEATURES  
Location/Qualifiers  
source  
1..4402  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"  
BASE COUNT 1259 a 1085 c 1111 g 947 t  
ORIGIN

alignment\_scores:  
Quality: 6095.50 Length: 1192  
Ratio: 5.174 Gaps: 1  
Percent Similarity: 98.826 Percent Identity: 96.980

alignment\_block:

US-09-171-553B-5 x AX052635  
Align seg 1/1 to: AX052635 from: 1 to: 4402

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
|||||  
224 GGAAGACGGGGTTCGACCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 273  
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
|||||  
274 GGAGGGCAACACAGTTGAGTTCTCTGGTTGATACCGAGCGGAACATTCAG 323  
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
|||||  
324 TGCTACTACAGCCATTAGGAAACTTAAAGATATAAAAAATCTGGGTGATG 373  
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67  
|||||  
374 GGTGCCACAGGGCAACACAGTATCCATGGACTACCCGGAAGAACAGTTGA 423  
67 leuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
|||||  
424 CTTGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGCC 473  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
|||||  
474 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGACAAATT 523  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
|||||  
524 TCTTTTGAACAAAGGAAACACAGAGTGTCTGCAATACACAAACCTATCAC 573  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
|||||  
574 TGTGTTGACCTCCAAATTAGATGAGATATATCGACTATATCTCTCCCTAG 623  
134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
|||||  
624 TAAAGCCTGATCAAAATATACAATTCGTGGTTGGAACAGTTTCCCAAGCC 673  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
|||||  
674 TGGGCAGAAACCCGAGGGATGGGTTGGCAAAAGCAAGTTCCCCCAAGT 723  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTrpProL 184  
|||||  
724 TATTCACCTGAAGGCCAGTGCACACAGTGTCTGAGTACAGACAGTACCCT 773  
184 euSerArgGluAlaArgGlyIleTrpProHisValGlnArgLeuIle 200  
|||||  
774 TGAGTAAAGAAAGCTCAAGAGGAATTCGGCGCATGTCCAAAGATTATTC 823  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
|||||  
824 CAACAGGCGCATCTTAGTCTCTGCTCAATCTCCCTGGAAATCTCCCTGCT 873  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
874 ACCGGTTAGAAGGCTGGGACTAATGACTATCGACACAGTACAGGACTTGA 923  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
924 GAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAACCT 973  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
|||||  
974 TATAACCTCTGTCTCTCCACCCCAACGAGAGCTGGTATACAGTATT 1023  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
|||||  
1024 GGACTTAAAGGATGCCCTTTTCTGCCCTGAGATTACACCCACCTAGCCAAC 1073  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300

||||| 1074 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 1123  
301 LeuThrTrpThrArgLeuLeuProGlnGlyPheLysAsnSerProThrIlePhe 317  
1124 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGACCACTCTT 1173  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
1174 TGACCAAGCCCTACACAGAGACCTGGCCCAACTTCAGGATCCACACCCCTC 1223  
334 InValThrLeuLeuGlnTyrValAspLeuLeuAlaGlyAlaThr 350  
1224 AGGTGACCCCTCCTCAGTACGTGGATGACCTGCTCTCGCGGAGGCCACC 1273  
351 LysGlnAspCysLeuGlnGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
1274 AAACAGGACTGCTTGAAGGACGCAAGGCACCTACTGCTGGAAATGTCGTA 1323  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
1324 CCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAGG 1373  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
1374 TAACATACTTGGGGTACAGTTTGGCGGACGCGGACGCGATGGCTGACGGAG 1423  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417  
1424 GCACGGAAGAAACCTGTAGTCCAGATACCGGCCCCACACACAGCCAAACA 1473  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
1474 AGTGAGAGAGTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 1523  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
1524 GGTTCGCGACCTTAGCAGCCCCACCTCTACCCACTAACCAAGAAAGAGG 1573  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
1574 GAATTCCTCGTGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 1623  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
1624 GGCCCTGTGTGAGCGCACCTGTCTGGCCCTCCCTGACGTAACTAAACCT 1673  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
1674 TTACCCCTTATGTGATGAGCGTAGGGAGTAGCCCCGGGGAGTTTAAACC 1723  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
1724 CAAACTCTAGGACCATGGAGAGACCTGTGCTACCTGTCAAAAGACT 1773  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
1774 CGATCCTGTAGCCAGTGGTGGCCGCTATGCTGAAGGCTATCGCAGCTG 1823  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
1824 TGCCCATACTGGTCAAGGACGCTGCAAAATGACTTTGGGACAGAAATA 1873  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567  
1874 ACTGTAATAGCCCCCATCGTGGAGACATCGTTCGCGACCCCCAGA 1923  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
1924 CCGATGGATGACCAACCGCCGATGACCCACTATCAAAAGCCCTCTCTCA 1973  
584 hClnArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600  
|||||

1974 CAGAGAGGGTCAGTTCGGCTCCACAGCCGCTCTCAACCTGCCACTCTT 2023  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
2024 CTGCTGAAGAGACTGATGAACAGTACTCATGATTGGCATCAACTATT 2073  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
2074 GATTGGAGGACTGGGGTCGCAAGGACCTTACAGACATACCGCTGACTG 2123  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
2124 GAGAAGTGTTAACCTGCTGCTCAGCAAGAGCAGCTATGTGGTGAAGGT 2173  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
2174 AAGAGGATGCTGGGGCGGGTGGTGGAGCGGACCCGACGATCTGGGC 2223  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
2224 CAGCAGCTCGCGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGCCC 2273  
684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700  
2274 TCAGCAGAGCTTTCGCGCTGGCCGAGGGAATCCATAACATTTATACA 2323  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
2324 GACAGCAGTATGCTTTGGGACTGCACACGTACACGGGGCCCATCTATA 2373  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
2374 GCAAGGGGGTGTCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 2423  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
2424 AAATTCCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 2473  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
2474 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGGAA 2523  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
2524 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCCGAGGGTGTAACTTC 2573  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
2574 TGCCTATGATAGAAACACCCAAAGCCCGGAGGAGCCGACGACAGTACACC 2623  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
2624 CTAGAAGACTGGCAAGAGATAAAGATAGACAGTCTCTGAGACTCC 2673  
817 oGluGlyThrCystThrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
2674 GGAAGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 2723  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
2724 AAGGTTAGATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 2773  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
2774 AAACACCTGCAGCAGTGGTGGTCAAACTCTCTTATCATGTTCTGAGGCT 2823  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
2824 ACCAGGAGTGGCTGATTGGTGGTCAAACTGTGTGCTCCCTGCCAGCTGG 2873  
884 aAsnAlaAsnProSerArgIleProGlyLysArgLeuLeuArgGlySer 900  
2874 TTATGTCTAATCCTTCCAGATACTCCAGAAAGAGACTTAAGGGGAAGC 2923



```
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
2924 CACCAGCGCTCACTGGAGTGGACTTCACGTAGGTAAAGCGGCTAA 2973
917 sYrGlyAsnLysTrpLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||
2974 ATACGGAACAAATATATCTATTTGGTTTGTAGACACCTTTTCAGATGG 3023
934 alGluAlaTyProThrLysLysGluThrSerThrValValAlaLysLys 950
|||||
3024 TAGAGGCTTATCTTAAAGAAAGACTTCAACCGTGGTGGCTAAGAA 3073
951 IleLeuGluLysPheProArgPheGlyIleProLysValLysLys 967
|||||
3074 ATACTGGAGAAATTTTCCAGATTGGATACCTAAGTATAGGTC 3123
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
|||||
3124 AGACAATGGTCCAGCTTTCGTTGCCAGGTAAATCAGGACTGGCCAAGA 3173
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyPheArgProGlnSerSer 1000
|||||
3174 TATTTGGGATGATTGGAACACTGCATTGTGCATACAGACCCCAAGCTCA 3223
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
|||||
3224 GGACAGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACTAATTT 3273
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034
|||||
3274 GACCGGGGAGCTGGGGTTAATGATTGGTAGCTCTCTGCGCTTTGTGC 3323
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrglu 1050
|||||
3324 TTTTGGGTTAGAACACCCCTGGAGAGTTGGGCTGACCCCTATGAA 3373
1051 LeuLeuTyrglyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
|||||
3374 TTACTCTACGGGGACCCCCCATTTGGTAGAAATGTCTTCGTACATAG 3423
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
|||||
3424 TGTGACGTGCTGCTTCCAGACCTTTGTTCTCTAGGCTCAAGGCACCTG 3473
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTySer... 1099
|||||
3474 ACTGGGTGAGACACAGCGCTGGAGCAACTCCGGGAGGCTTACTCAGGA 3523
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
|||||
3524 GGAGGAGACTTGCAGATCCACATCGTTTCCAAGTGGGAGATTCAGTCTA 3573
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
|||||
3574 CGTATAGCCGCCCGTGCAGGAAACCTTCAGACTCGGTGGAAGGCCCTT 3623
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
|||||
3624 ATCAGCTACTTTTGACCACACACCGGCTGTGAAAGTGAAGAAATCTCC 3673
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166
|||||
3674 ACTTGGATCCATGCATCCACCTTAAGCGCGGCCACCTCCCGATTCCGG 3723
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
|||||
3724 GTGGAAGCCGAAAGACTGAAATTCCTTAAAGCTTCGCTCCATCCGCG 3773
1183 euValProTyrgSerAsnAsnSer 1191
|||||
3774 TGGTTCCTTACTCTGCAATACTCC 3799
```

```
seq_name: gb_pat:AX052634
seq_documentation_block: 7362 bp DNA PAT 12-JAN-2001
LOCUS AX052634
DEFINITION Sequence 19 from Patent WO0071726.
ACCESSION AX052634
VERSION AX052634.1 GI:12226824
KEYWORDS
SOURCE
  ORGANISM
    porcine endogenous retrovirus.
  Viruses; Retrovirus; Retroviridae; Mammalian type C
  retroviruses; 1-Mammalian type C virus group.
  1 (bases 1 to 7362)
  AUTHORS
    Federspiel, M.J.
  TITLE
    Methods to inhibit infectious agent transmission during
    xenotransplantation
  JOURNAL
    Patent: WO 0071726-A 19 30-NOV-2000;
    Mayo Medical Ventures (US)
  FEATURES
    Location/Qualifiers
      1..7362
        /organism="porcine endogenous retrovirus"
        /db_xref="taxon:61673"
BASE COUNT 1997 a 1821 c 1881 g 1663 t
ORIGIN
alignment_scores:
  Quality: 6095.50 Length: 1192
  Ratio: 5.174 Gaps: 1
  Percent Similarity: 98.826 Percent Identity: 96.980
alignment_block:
  US-09-171-553B-5 x AX052634 ..
  Align seg 1/1 to: AX052634 from: 1 to: 7362
1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2823 GGAAGACGGGTTCGGACCCCTCCCGAGCCGAGGTAACTTTGAAGGT 2872
17 iGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2873 GGAGGGCAACACAGTTGAGTTCTCTGGTTGATACCGGAGCGAAACATTCA 2922
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2923 TGCCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 2972
51 GlyAlaThrGlyGlnArgGlnTyProTrpThrArgArgThrValAs 67
|||||
2973 GGTGCCACAGGCAACACAGTATCCATGGACTACCCGAGAACAGTTGA 3022
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
3023 CTTGGAGTGGGAGGGTAAACCCACTGTTTCTGTCTCATACCTGAGTGCC 3072
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
3073 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGACAAAT 3122
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
3123 TCTTTTGAACAAGGAAACACAGAGTGTCTGCAATAACAAACCTATCAC 3172
117 rValLeuThrLeuGlnLeuAspGlyTyArgLeuTyArgLeuTySerProGlnV 134
|||||
3173 TGTCTTGACCTCCCAATTAGATGACGAATATCGACTATACTCTCCCTAG 3222
134 alLysProAspGlnAspIleGlnSerTrpLeuGlnPheProGlnAla 150
|||||
3223 TAAAGCCTGATCAAAATATACAAATCTGGTTGGAAACAGTTTCCCAAGCC 3272
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
```

||||| 3273 TGGCAGAAACCGCAGGATGGTTGGCAAGCAAGTTCCTCCACCAAGT 3322  
167 lileGlnLeuLysAlaSerAlaThrProValSerValArgGlnThrProL 184  
3323 TATTCACCTGAAGGCCAGTGCACACACAGTGTCACTGACAGTACCCCT 3372  
184 euserArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
3373 TCAGTAAAGAGCTCAAGAGGAATTCGGCCGCATGTCCAAAGATTATC 3422  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
3423 CAACAGGCATCCTAGTTCCTGTCCAATCTCCCTGGATATCTCCCTGCT 3472  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
3473 ACCGGTTAAGAACTGGGACTAATGACTATCGACCATGACAGACTTGA 3522  
234 tGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
3523 GAGAGTCAATAAAGCGGTGCAGGATATACACCAACAGTCCCGAACCT 3572  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
3573 TATAACCTCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATT 3622  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
3623 GGACTTAAAGGATGCTTTTCTGCTCGAGATTACACCCCACTAGCCCAAC 3672  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
3673 CACTTTTTCCTTCAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3722  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
3723 CTCACCTGGACCCGACTGCCCAAGGGTTCAGAACTCCCGGACCATCTT 3772  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3773 TGACGAAGCCCTPACACAGACCTCGCCCACTCAGGATCAACACCCCTC 3822  
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
3823 AGGTGACCTCTCCAGTACGTGGATGACTGCTTCTGGCGGAGCCACC 3872  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
3873 AAACAGGACTGCTTAAAGGCCACGAAGGCCTACTGCTGGAATTGCTGA 3922  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3923 CTTAGCTACAGAGCTCTGCTTAAAGCCCAAGATTCAGGAGAGAGG 3972  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
3973 TAACTACTTGGGTACAGTTTTCGGGAGCGGCAGCGATGGCTGACGGAG 4022  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
4023 GCACGGAAGAAACTGTAGTCCAGATACGGCCCAACACAGCCAAACA 4072  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
4073 AGTGAGAGAGTTCCTGGGACAGCTGGATTTCAGAGCTGTGATCCCGG 4122  
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlnLysGly 450  
4123 GGTTCGACCTTAGCAGCCCACTCTACCCACTAACCAAGAAAAAGGG 4172  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysL 467  
|||||

4173 GAATTCCTCTGGCTCCTGAGCACACAGGCAATTTGATGCTATCAAAAA 4222  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484  
4223 GGCCTTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCT 4272  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
4273 TTACCTTTATGTGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 4322  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
4323 CAACCTCTAGGACCAATGGAGAGACCTGTTCCTACTCTCAAGAGAGCT 4372  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534  
4373 CGATCTGTAGCCAGTGGTTGGCCCGTATGCCCTGAAGGCTATCGCAGCTG 4422  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
4423 TGGCCATACTGGTCAAGGAGCGTGACAAATTTGACTTTGGGACAGATAATA 4472  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
4473 ACTGTATAGCCCTCCATGCGTTGGAGACATCGTTCCGCGAGCCCCAGA 4522  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
4523 CCGATGATGACCAACGCCCGCATGACCCCACTATCAAAAGCTGCTTCTCA 4572  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
4573 CAGAGAGGGTCAGTTGCTCCACCGCCCTCTCAACCTGCCACTCTT 4622  
601 LeuProGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
4623 CTGCTGAAGAGCTGATGAACCACTGACTCATGATGCCATCACTATT 4672  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4673 GATTGAGGAGACTGGGTCGCCAAGGACCTTACAGACATACCGCTGACTG 4722  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
4723 GAGAAGTGTAACTTGGCTTCACTGACGGAAGCAGCTATGTGGTGGAGGT 4772  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4773 AAGAGGATGCTGGGCGCGCGTGGTGGACGGGACCGCCGACGATCTGGGC 4822  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684  
4823 CAGCAGCTGCCGGAAGAACTTTCAGACAAAAGGCTGAGCTCATGGCCC 4872  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4873 TCAGCAGAGCTTTGCGGCTGCCGGAAGGAATCCATAACATTATATA 4922  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
4923 GACAGCAGGTATGCTTTGCGACTGTCACAGCTACACGGGGCCATCTATAA 4972  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4973 GCAAGGGGGTGTCTTACCTCAGCAGGGGGAATAAAGAACAAAGAGG 5022  
734 luIleLeuSerLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
5023 AAATTCATAGCCTATTAGAGCCTTACATTTGCCAAAAAGGCTAGCTATT 5072  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuLeuSerArgGlyAs 767  
5073 ATACACTGCTCTGACATCAGAAAGCCAAAGATCCCATATATCCAGAGGAA 5122  
|||||

767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAlaSerLeuL 784  
|||||  
5123 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCCGAGGGTTAAACCTTC 5172  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
5173 TGCCTATGATAGAAACACCCAAAGCCCCAGAACCCCGGACGACAGTACACC 5222  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5223 CTAGAACACTGGCAAGAGATAAAGATAGACAGTCTCTGAGAGTCC 5272  
817 oGluGluThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
5273 GGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCTCCGCCACANAG 5322  
834 LuGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
5323 AAGGGTTAGAAATATGTCCACAGATACATCGTCTAACCCACCTAGGAAC 5372  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
5373 AAACACCTGCAGCAGTTGGTCAGAACATCTCTTATCATGTTCTGAGGCT 5422  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
5423 ACCAGGAGTGGCTGATTCGGTGGTCAACACTGTGTGCCCTGCCAGCTGG 5472  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
5473 TTAATGCTAATCTCTCCAGAAATACCTCCAGGAAAGAGACTAAGGGGAAG 5522  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
5523 CACCCAGGCGCTCACTGGGAAGTGGACTTCACCTCAGGTAAAGCCGGCTAA 5572  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
5573 ATACGGAACAATACTATTTGTTGTAGACACCTTTTCAGGATGGG 5622  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
5623 TAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 5672  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
5673 ATACTGGAGGAATTTTCCAGATTGGGAATACCTAAGGTAATAGGGTC 5722  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5723 AGACAATGGTCCAGCTTCCTGCCAGGTAAAGTCAAGGACTGGCCCAAG 5772  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
5773 TATTTGGGATTTGATGGAACTCATTTGTGCATACAGACCCCAAGCTCA 5822  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
5823 GGACAGGTAGAGAGATGAATAGAACCATTTAAAGAGACCTTACTAAAT 5872  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValI 1034  
|||||  
5873 GACCGGGAGACTGGGGTTAATGATTTGGATAGCTCTCTGCCCTTTGTAC 5922  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
5923 TTTTATAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5972  
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5973 TTACTACGGGGGACCCCGCCCATTTGGTAGAAATTTGCTTCGTACATAG 6022

1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
6023 TGCTGACGTGCTGCTTTCCAGCCTTTGTCTTAGGCTCAAGGCACCTTG 6072  
1084 LuTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
|||||  
6073 AGTGGGTGAGAACACGAGCGTGGAGGCAACTCCGGGAGCGCTACTCAGGA 6122  
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116  
|||||  
6123 GGAGGAGACTTGCAGATCCACATCGTTTCCAAGTGGGAGATTCAGTCTA 6172  
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProF 1133  
|||||  
6173 CGTTAGACGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGCCCTT 6222  
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
|||||  
6223 ATCAGCTACTTTTGACCACCAACGCGCTGAAAGTCGAAGGAATCTCC 6272  
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGl 1166  
|||||  
6273 ACCTGGATCCATCCATCCACGTTAAGCGCGGCCACCTCCCGATTCCGG 6322  
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
|||||  
6323 GTGGAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 6372  
1183 euValProTyrSerAsnAsnSer 1191  
:::|||||  
6373 TGGTCTTACTCTGTCAATAACTCC 6398

seq\_name: gb\_om:AF038600

seq\_documentation\_block:

LOCUS AF038600 8132 bp mRNA MAM 21-MAY-1998  
DEFINITION Sus scrofa porcine endogenous retrovirus PERV-MSL, complete  
sequence.  
ACCESSION AF038600  
VERSION AF038600.1 GI:3133301  
KEYWORDS  
SOURCE  
ORGANISM  
pig.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE  
AUTHORS  
Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and  
Fishman,J.A.  
IDENTIFICATION of a full-length cDNA for an endogenous retrovirus  
of miniature swine  
J. Virol. 72 (5), 4503-4507 (1998)  
98216827  
REFERENCE  
2 (bases 1 to 8132)  
AUTHORS  
Fishman,J.A.  
TITLE  
Direct Submission  
Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts  
General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
REFERENCE  
3 (bases 1 to 8132)  
AUTHORS  
Fishman,J.A.  
TITLE  
Direct Submission  
Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts  
General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
REMARK  
Sequence update by submitter  
COMMENT  
On May 16, 1998 this sequence version replaced gi:3116441.  
FEATURES  
source  
1..8132  
/organism="Sus scrofa"  
/strain="miniature swine"  
/db\_xref="taxon:9823"  
/cell\_type="lymphocytes"  
/note="porcine endogenous retrovirus PERV-MSL"  
1..564  
585..2159

LTR  
CDS

```

/codon_start=1
/product="gag protein"
/protein_id="AAC16763.1"
/db_xref="GI:3116442"
/translation="MGQVTTPSLTLDHWTVEKSAHNLVQVKKGPMQWTCFVSEMP
TFDVGPSGTFNSEIILAVKAVIFOTGCGSHDQEPYLLTWODLAEDPPVWVKFWLN
KPKKPRILALGEKKNKSAEKVSPHPIEIEBPAMPPEOSPPPPYLAQGAFFS
FPAPGAPAVEGPAAGTSGRATPERTDEIATPLRTYPTPGQLOLPQWPFSS
ADLYNKKTHNPESDQRLTGLVSLMFHQPTWDDCOQLQTLFTYEEERILLLEA
KRVNCGADGPRTRLONEIDMGRLPTRGWDYNTAEGRESLKIVROALVAGERRR
TNLAKVREYMQGNEPSPVLERLLEAFRYVTPDPTSEAQKASVALAFIGOSALDIR
KKQRLLEGLOAEKRLDLVKAEEKYIKRETEERQKREEREEREERNKRNKQENLIT
KILAAVVEKSNTEREDFRKINSRQSGNLGNRTPLDKDQCAVCKERKGRHWARNCPK
KGNKGPRILALEEDKD"
2307..5744

```

CDS

```

/codon_start=1
/product="pol protein"
/protein_id="AAC16764.1"
/db_xref="GI:3133302"
/translation="MGATGQQQYPTWTRRTVDLGVGRVTHSLVIEPCAPLLGRDLL
TKMAQISFEQKPEYSANNKPTVLTLDDEYRLYSPLVKPDONIQFWLEQFPQAW
ATFAGAGLAKQVPOVIOLKASATPVSRQYPLSKAEQEGIRPHVORLIQOQGIIPVQ
SPWNTPLLPVRKPGTNDYRPVQDLREYNKRVQDIHTYVNPYNLICALPQPSWTVL
DLKDAFFLRHLPTSQPLFAEWRDEGTGOLTWTRUPQGGKNSPTIFDDEALHRLD
ANFRHQPTVLLQYVDDLLLAGATKQDCLEGTALLLELSDLGYSASAKKAOICRE
VYLYGLSRDGGRLWTEARKTVVQIPAPTTAKQREFLTAGFCRLWTPGFAALAP
LYLTKEGEFSWAPHEKAFDAIKKALLSAPALAPDVTKPTFLYVDRKGVARGVL
TOPLGPRRPVAVLSKLLDPVAGSWPICLKATAAVAILVKDADKLTGONITVIAPHA
LENIVRQPDRTWNTNARMTHYQSLLLTERTVFPAPPAALNPATLLPEETDEPVTDOHO
LLEETGPRKDLTDIPLTGEVLTFTDGSYVVEGRMAGAAVVDGTRTWASSLPGG
TSAQKALMALQALRAEKSINIYTDRIYAFATAHVHGAIIYKORGLTLAGREIKN
KEBILSLAEALHPLKRLAIHCPGHOKAKDLISRGNMADRAVAKAQAQVNLPIIEM
PKAPEPRQVTLDEWQEIKKIDQFSETPEGTCTYSDGKEILPHKEGLEVYQIHLTH
LGTKHLQQLVRTSPYHVLRLPGVADSVVVKHCVPCOLVANPSPMPKRLGRSHPGAH
WEYDFTEVPKAKGNKYLIVFVDTSEGVVEAYPTKETSTVVAKKILFIFPRFGIPK
VIGSDNGPAPFVHVSQGLAKILGIDWKHLCAVTRPQSSGOVERMNRITKELLKLAET
GNDWMLALPLFVLVRNTPGQPLTYELLYGGPPPLVEIASVHSADVLTLSPQFSR
LKALEWRQARQRLREAYSGGGDLQIPHRFQVGSVYVRRHAGNLFTFRWKGPLYVL
LTPYAVKVEGISTWIHASHVRPAPPDPSGMAKENTENPLKRLHRVVPYSVYNLSD"
5620..7536

```

CDS

```

/codon_start=1
/product="env protein"
/protein_id="AAC16765.1"
/db_xref="GI:3133303"
/translation="MHPTLNRRHLPIRGKPKRLKIPLSFASITAMFLTSLTSITQNGM
RIGDSLNSHKPLSLTLTDSGTGINNTQGEAPLGTWMPDLYLCRLSVIPSLTSP
DILHANGFYVCPGPNNGKHCNPRDFCKQWNCVTSDNGYKWPSTSOODRYSFYSVN
TYTSQGFNLTWIRGSPKSPSDDLKLSFTKQKQENILKWNGSMGVMVYGG
SGKQPSLIRUKINOLEPPMAIGNTVLTGQRPPTQGPFSNNITSSDPTESST
TRMGAKLSLIQAFQALNSTTPEATSSCMLCLASGPPYYEGMARGFNVTKERDQ
CTWGSQNKLTLEVSGKTCIGKVPSSHQHLNHTFAFNQTSOYLVPYDRWACN
TGLTPCVSTLVNOTKDFCIMOIVPRVYYPEKALIDEYDYNHRKREKREISLTIAV
MLGLGVAAGVTGTAALVTGPOLETGLSNLHRIVTEDLQALEKSVSNLESLTISE
VVLONRRGLDLLFLKEGGLCVALKKECCFYVDHSGAIRDSMNKRLRLEKRRERET
QGFEGFNRLSLATLALLSALTGLPLVLLLLTLTVGPCIINKLFIPIREKISAVQIMVL
ROOYQSPSSREAGR"
7518..7976

```

LTR

BASE COUNT 2247 a 1977 c 2037 g 1871 t

ORIGIN

## alignment\_scores:

Quality: 6087.50 Length: 1191  
 Ratio: 5.168 Gaps: 1  
 Percent Similarity: 98.908 Percent Identity: 96.725

## alignment\_block:

US-09-171-553b-5 x AF038600

Align seg 1/1 to: AF038600 from: 1 to: 8132

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

|||||

```

2160 GGGAGAGCGGGTTCGGACCCCTCCCGAGCCCGAGGGTAACTTTGAAGGT 2209
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2210 GGAGGGCAACCAAGTTGAGTTCTCTGTTGATACCGGAGCGAAACAATCAG 2259
34 allLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2260 TGTACTACAGCCATTAGGAAAACTAAAAAGATAAAAAATCCTGGGTGATG 2309
51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgArgThrValAs 67
|||||
2310 GGTGCCACAGGGCAACAAGTAGTATCCATGAGCATACCCGAGAGAACAGTTGA 2359
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
2360 CTTGGGAGTGGGACGGGTAACCCACCTCGTTTCTGGTCATACCTGAGTGCC 2409
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2459
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2460 TCTTTTGAACAAGGGAACCCAGGAAGTGTCTGCAAAATAACAACCTATCAC 2509
117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
|||||
2510 TGTGTTGACCTTCCAAATTAGATGACGAATATGACTATATCTCTCCCTAG 2559
134 allLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2560 TAAAGCCTGTATCAAAATATACAAATCTCTGTTGGAACAGTTTCCCCAGCC 2609
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2610 TGGCAGAACCCCGAGGATGGTGTGGCAACCAAGTTCCCCCAAGT 2659
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2660 TATTCAACTGAAGGCCAGTCCACACAGTGTCTCAGTCAGACAGTACCCCT 2709
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
|||||
2710 TGAGTAAGAAGCTCAAGAAAGGAATTCGGCCGATGTCCAAAGATTAACT 2759
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
2760 CAACAGGCCATCTAGTCTCTGTCCTCAATCTCCCTGGAATCTCCCTGCT 2809
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2810 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCATGACAGGACTTGA 2859
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||
2860 GAGAGGTCATAAACGGGTGAGGATATACCCCAACAGTCCCGAACCCCT 2909
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
|||||
2910 TATAACCTCTTGTGCTCTCCCAACCCAGGAGCTGGGTATACAGTATT 2959
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2960 GGACTTAAAGGATGCTCTCTCTGCTGAGATTACACCCCACTAGGCAAC 3009
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3010 CACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCCGGCAG 3059
301 LeuThrTrpThrArgLeuProGlnGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3060 CTCACCTGGACCGGACTGCCCCAGGGGTTCAGAACTCCCCCGACCATCTT 3109

```

```

317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3110 TGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCCCTC 3159
|||||
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3160 AGGTGACCCCTCTCCAGTAGTGATGACCTGCTTCTGGGGAGGCCACC 3209
|||||
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
|||||
3210 AAACAGAGACTGCTTAGAAGGCACCAAGGCACACTACTGCTGGAATTTGTCGA 3259
|||||
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3260 CCTAGGCTACAGAGCCCTCTGCTAAGAAGGCCAGATTTCAGGAGAGAGG 3309
|||||
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrrpLeuThrGlu 400
|||||
3310 TAACATACCTTGGGGTACAGTTTACGGGACGGGAGCGATGCTCACGGAG 3359
|||||
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG1 417
|||||
3360 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACAGCCAAACA 3409
|||||
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3410 AATGAGAGAGTTTTTGGGGACAGCTGGATTTCGACAGCTGTGGATCCCGG 3459
|||||
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3460 GGTTCGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGG 3509
|||||
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3510 GAATTCCTCTGGGCTCTGAGCACCAAGAGGCATTTGATGCTATCAAAA 3559
|||||
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3560 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAACCCCT 3609
|||||
484 hethrLeuTyrValAspGluArgGlyGlyValAlaArgGlyValLeuThr 500
|||||
3610 TTACCCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGGAGTTTTAAC 3659
|||||
501 GlnThrLeuGlyProTrpArgArgProValAlaIleLysLysLe 517
|||||
3660 CAAACCCCTAGACCATGGAGAAGACCTGTGCGCTACCTGTCAAAGAAAGCT 3709
|||||
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
|||||
3710 CGATCTCTGTAGCCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759
|||||
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
|||||
3760 TGGCCATCTACTGGTCAAGGAGCGCTGACAAAATTGACTTTGGACAGAAATA 3809
|||||
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProAs 567
|||||
3810 ACTGTATATACCCCATGATTTGGAGAACATCGTTTCGGCAGCCCCCA 3859
|||||
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
|||||
3860 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCCTCTCTCA 3909
|||||
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
|||||
3910 CAGAGAGGGTCAAGTTCGCTCCACCAGCGCTCTCAACCCCTGCCACTCTT 3959
|||||
601 LeuProGluThrThrAspGluProValThrHisAspCysHisGlnLeuLe 617
|||||
3960 CTGCCCTGAAGAGACTGATGAACACAGTGACTCATGATTGCCATCAACTATT 4009
|||||

```

```

617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
|||||
4010 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4059
|||||
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
|||||
4060 GAGAAGTGCTAACTGTTCACTGACGGAAGCAGCTATGTGTGGGAAGGT 4109
|||||
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
|||||
4110 AAGAGGATGGCTGGGGCGGGTGGTGACGGGACCCGCCACGATCTGGGC 4159
|||||
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
|||||
4160 CAGCACCCCTGCCGGGAGGAACTTCAGCACAAAAGGCTGAGCTCATGGCCC 4209
|||||
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
|||||
4210 TCACGCAAGCTTTGGCGCTGGCCGAAGGGAATCCATAAACATTTATAGC 4259
|||||
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
|||||
4260 GACAGCAGGTATGCCCTTTCGACTGCACACGTACATGGGGCCATCTATAA 4309
|||||
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
4310 ACAAGGGGGTTCCTTACCTCAGCAGGAGGGAATAAAGACAAGAGG 4359
|||||
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
4360 AAATTCTAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTATT 4409
|||||
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
4410 ATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAA 4459
|||||
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784
|||||
4460 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCCCGGCGGTGTTAACCTTC 4509
|||||
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
4510 TGCCCTATAATAGAAATGCCAAAGCCCAAGAACCCAGACGACAGTACACC 4559
|||||
801 LeuGluAspTrpGlnGluIleLysLysIleaspGlnPheSerGluThrPr 817
|||||
4560 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTCTGAGACTCC 4609
|||||
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
4610 GGAAGGGACCTGCTATACCTCAGATGGGAAGGAATCTGCCCCACAAG 4659
|||||
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
4660 AAGGGTTAGATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 4709
|||||
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
4710 AAACACCTGCGACGAGTGTGTCAAGACATCCCTTATCATGTTCTGAGGCT 4759
|||||
867 uProGlyValAlaAspSerValLysHisCysValProCysGlnLeuV 884
|||||
4760 ACCAGAGTGGCTGACTCGGTGGTCAACATTTGTGTGCCCTGCCAGCTGG 4809
|||||
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4810 TTAATGCTAACTCTCCAGAAATGCCGCCAGGGAAGAGACTAAGGGGAAGC 4859
|||||
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4860 CACCAGGCGCTCACTGGGAAGTGGACTTCACGTAGGTAAAGCCGCTAA 4909
|||||
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
|||||

```

```
|||||
4910 ATACGGAACAACAAATACCTATTGGTTTGTAGACACCTTTTCAGATGGG 4959
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
4960 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAGAA 5009
951 ileLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967
5010 ATACTGGGAAGAAATTTTCCAAAGATTGGAATACCTAAGGTAATAGGTC 5059
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
5060 AGACAATGCTCAGCTTTTGTGGCCACGTAAGTCAGGACTGCCCAAGA 5109
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
5110 TATTGGGATGATTGGAAACTGCAATGTGCATACAGACCCCAAGCTCA 5159
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLysLe 1017
5160 GGACAGGTAGAGAGATCAATGAACCACTTAAGAGACCCCTTACTAAAT 5209
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
5210 GACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTGTGC 5259
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
5260 TTTTATAGGTTAGGAACACCCCTGGACAGCTTTGGCTGACCCCTATGAA 5309
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
5310 TTACTCTACGGGGACCCGCCCATTTGTTAGAAATGCTTCTGTACATAG 5359
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
5360 TGCTGATGTGCTGCTTCCAGCCCTTTGTTCTTAGGCTCAAGGCACCTG 5409
1084 lutrPValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
5410 AGTGGGTGAGACAACGCGTGGAGGCACTCCGGAGGCCCTACTCAGGA 5459
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
5460 GGAGGAGACTTGCAGATCCCACTCGTTTCCAAAGTGGGAGATTTCAGTCTA 5509
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
5510 CGTTAGACGCCACCGTGGAGAAACCTCGAGACTCGGTGGAGGCCCTT 5559
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
5560 ATCTGCTACTTTTGACACACACCGCTGTGAAAGTCGAAGGAATCTCC 5609
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166
5610 ACCTGGATCCATCATCCACCGCTTAAACCGCGCCACCTCCGATTCGGG 5659
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
5660 GTGGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCG 5709
1183 euValProTyrTrpSerAsnAsn 1190
5710 TGGTTCCTTACTCTGTCAATAAC 5732
seq_name: gb_om:AF038599
seq_documentation_block:
LOCUS AF038599 8132 bp mRNA MAM 21-MAY-1998
DEFINITION Sus scrofa porcine endogenous retrovirus Tsukuba-1 mRNA, complete
sequence.
```

LTR 7548..8132  
 BASE COUNT 2247 a 1980 c 2030 g 1875 t  
 ORIGIN

alignment\_scores:  
 Quality: 5979.50 Length: 1193  
 Ratio: 5.076 Gaps: 3  
 Percent Similarity: 98.743 Percent Identity: 96.479

alignment\_block:

US-09-171-553B-5 x AF038599 ..

Align seg 1/1 to: AF038599 from: 1 to: 8132

```

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
2168 GGGAGAGCGGGTTCGGACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGT 2217
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
2218 GGAGGGCAACCAAGTTGAGTTCTGCTGATACCGGAGGAAACATTGAG 2267
34 alleuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
|||||
2268 TGTACTACTACGCCATTAGGAAACTAAAGATAAAAAATCTCGGTGATG 2317
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgGlnThrValAs 67
|||||
2318 GTTGC.ACAGGGCAACACAGTATCCATGGACTACCCGAAG.ACAGTTGA 2365
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
|||||
2366 CTTGGGAGTGGGACGGTAAACCCACATCGTTCTGCTCATACCTGAGTGC 2415
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
2416 CAGCACCCCTCTTAGGTAGACATTAATGACCAAGATGGGAGCACAAAT 2465
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
2466 TCTTTTGACACAGGGAACAGAGTGTCTGCAAAATAACAACTATCAC 2515
117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134
|||||
2516 TGTGTTGACCCCTCCAATTAGATGACCAATATCGACTATACTCTCCCTAG 2565
134 allysProaspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
2566 TAAAGCCTGATCAAAATATACAAATCTGTTGGTGGACAGTTTCCCAAGCC 2615
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
|||||
2616 TGGGCAGAACCCGAGGATGGTTGGCAAGCAAGTTCCCCCAAGT 2665
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
|||||
2666 TATTCAACTGAAGCCAGTGCACACACAGTGTGAGTCAGTCAGACAGTACCCT 2715
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeulle 200
|||||
2716 TGAGTAAGAAGCTCAAGAAGAAATTCGGCCGCATGTCCAAAGNTAATC 2765
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
|||||
2766 CAACAGGGATCCTAGTTCTCTGCCAATCTCCCTGGAATATCTCCCTGCT 2815
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
|||||
2816 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACAGTACAGGACTTGA 2865
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
|||||

```

```

2866 GAGAGGTCAATAAACGGTGCAGGATATACACCAACAGTCCCGAACCT 2915
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValle 267
|||||
2916 TATAACCTCTTGTGTCTCTCCACCCCAACGGAGCTGGTATACAGTATT 2965
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
|||||
2966 GGACTTAAAGGATGCTTCTCTGCTGAGATTACACCCACTAGCCAC 3015
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
|||||
3016 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3065
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
|||||
3066 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3115
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3116 TGACGAAGCCCTACACAGAGACTGGCCAACTTCAGGATCCACACCCCTC 3165
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3166 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3215
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367
|||||
3216 AAACAGGACTGCTTAGAAGGCACGAAGGCACCTACTGCTGGAATTGCTGA 3265
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3266 CCTAGGCTACAGAGCTCTGTAGAACGCCAGATTTGCAGAGGAGAGG 3315
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
|||||
3316 TAACATACTTGGGTACAGTTTTACGGGACGGCAGCGATGGCTGACGGAG 3365
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417
|||||
3366 GCACGGAGAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3415
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3416 AATGAGAGAGTTTTGGGACAGCTGGATTTGCAGACTGTGGATCCGG 3465
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlyLysGly 450
|||||
3466 GGTTCGACCTTAGCAGCCCACTCTACCCGCTAACCCAAAGAAAAAGGG 3515
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||||
3516 GAATTCCTCTGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAAA 3565
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
|||||
3566 GGCCCTGCTGAGCGACCTGCTGGCCCTCCCTGACGTAACTAAACCT 3615
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
|||||
3616 TTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 3665
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
|||||
3666 CAACCCCTAGGACCATGGAGAGACCTGTGCTGCTTACCTGTCAAAAGAGCT 3715
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534
|||||
3716 CGATCTCTGACCCAGTGGTTGGCCCATATGCTCAAGGCTATCCAGCTG 3765
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGln.AsnIl 550
|||||
3766 TGGCATACTGCTCAAGGACGCTGACAAATTTGACTTTGGGACAAAGAAAT 3815

```



550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProA 567  
|||||  
3816 AACTGTAATAGCCCCCATGATGGAGAACATCGTTCGGGAGCCCCAG 3865  
|||||  
567 sPArgTrpMetThrasnAlaArgMetThrHisTyrGlnSerLeuLeu 583  
|||||  
3866 ACCGATGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTTCTC 3915  
|||||  
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600  
|||||  
3916 ACAGAGAGGGTACGTCCTGCCACCAACCGCTCTCAACCCCTGCCACTCT 3965  
|||||  
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617  
|||||  
3966 TCTGCTCAAGAGACTGATGAACAGTGCATCATGATTGCCATCAACTAT 4015  
|||||  
617 euIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633  
|||||  
4016 TGATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACT 4065  
|||||  
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluL 650  
|||||  
4066 GGAGAAGTGTAACTGGTTCACCTGCCGAGACGACTATGTGGTGGAAAG 4115  
|||||  
650 yLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpA 667  
|||||  
4116 TAAGAGGATGGTGGGCGCGGTGGTGACGGGACCGCGCATCTGGG 4165  
|||||  
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683  
|||||  
4166 CCAGCAGCGCTCGCGGAGGAGAACTTCAGCACAAAGGCTGAGCTCATGGCC 4215  
|||||  
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrTh 700  
|||||  
4216 CTCACGCAAGCTTCGCGCTGCGCAAGGGAATCCATAAACATTTATAC 4265  
|||||  
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrL 717  
|||||  
4266 GGACAGCAGGTATGCTTTGGCAGTGCACACGATACATGGGCGCATCTATA 4315  
|||||  
717 ySlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGlu 733  
|||||  
4316 AACAAAGGGGTGCTTACCTCAGCAGGGGAGGAAATAAGAACAAAGAG 4365  
|||||  
734 GluIleLeuSerLeuGluAlaLeuHisLeuProLysArgLeuAlaI 750  
|||||  
4366 GAAATCTAAGCCTATTAGAAGCCGTACATTTACCAAAAGGCTAGCTAT 4415  
|||||  
750 eIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767  
|||||  
4416 TATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAA 4465  
|||||  
767 sGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeu 783  
|||||  
4466 ACCAGATGCTGACCGGGTGCACAGCAGGCGCCAGCCAGGGGTAAACCTT 4515  
|||||  
784 LeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrTh 800  
|||||  
4516 CTGCTATTAATAAGAAATGCCAAAGCCCGCCAGAACCCAGACAGCATAC 4565  
|||||  
800 rLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrP 817  
|||||  
4566 CCTAGAAGACTGGCAAGAGATAAAAGATAGACCA.TTCTCTGAGACTC 4614  
|||||  
817 roGluGlyThrCysThrThrSerTyrGlyLysGluIleLeuProHisLys 833  
|||||  
4615 CGGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTTCGCCCCACAAA 4664  
|||||  
834 GluGlyLeuGluTyrValGlnGln.IleHisArgLeuThrHisLeuGlyT 850  
|||||  
4665 GAAGGTTAGANATATGTCACCAAGATACATGCTTAACCCACCTAGGAA 4714  
|||||  
850 hrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArg 866  
|||||  
4715 CTAACACCTGCAGCAGTGGTTCAGAACATCCCTTATCATGTCTGAGG 4764  
|||||  
867 LeuProGlyValAlaAspSerValLysHisCysValProCysGlnLe 883  
|||||  
4765 CTACCAGAGTGGCTGACTCGGTGGTCAACATTTGTGCGCTGCCAGCT 4814  
|||||  
883 uValAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlyS 900  
|||||  
4815 GGTATATGTAATCTTCAGATGCTCCAGGAAGAGACTAAGGGGAA 4864  
|||||  
900 erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAla 916  
|||||  
4865 GCCACCCAGCGCTCACTGGGAAAGTGCATTCACCTGAGGTAAGCGCGCT 4914  
|||||  
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933  
|||||  
4915 AAATATGGAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGAGTG 4964  
|||||  
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950  
|||||  
4965 GGTAGAGGCTTATCTTAAAGAAAGAGACTTCAACCGTGGTAGCTAAAA 5014  
|||||  
950 ySlnLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966  
|||||  
5015 AAATACTGGAAGAAATTTTCCAAGATTTTGAATACCTAAGGTAATAGG 5064  
|||||  
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaL 983  
|||||  
5065 TCAGACAAATGGTCCAGCTTTTGTGCCAGGTAACTCAGGACTGGCCAA 5114  
|||||  
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS 1000  
|||||  
5115 GATATTGGGATGATTGGAACTGCATTTGTCATACAGACCCCAAGCT 5164  
|||||  
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016  
|||||  
5165 CAGGACAGGTAGAGAGGATGAATAGAACCATTAAGAGAGACCCCTTACTAA 5214  
|||||  
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033  
|||||  
5215 TTGACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCTGCCCTTGT 5264  
|||||  
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050  
|||||  
5265 GCITTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATG 5314  
|||||  
1050 luLeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHis 1066  
|||||  
5315 AATTACTCTACGGGGGACCCCCCATTTGGTAGAATTTGCTCTGTACAT 5364  
|||||  
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083  
|||||  
5365 AGTCTGATGCTGCTCTTCCAGCCTTTGTTCTCTAGGCTCAAGCACT 5414  
|||||  
1083 uGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099  
|||||  
5415 TGAGTGGGTGAGACACAGCGTGGAGCAACTCCGGGAGGCGCTACTCAG 5464  
|||||  
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115  
|||||  
5465 GAGGAGAGACTTGCAGATCCCATCGTTTCCAAGTGGGAGATTACGTC 5514  
|||||  
1116 TyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyPr 1132  
|||||  
5515 TACGTTAGACGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGCCCC 5564  
|||||  
1132 oTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIleP 1149  
|||||  
5565 TTATCTCGTACTTTTGACCACACACCGCTGTGAAAGTCCAAAGGAATCT 5614  
|||||  
1149 roThrTrpIleHisAlaSerHisValLysProAlaProProProAspSer 1165  
|||||

|||||  
 5615 CCACCTGGATCCATCCACCTTAAGCGCGCCACCTCCCGATTCCG 5664  
 1166 GlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisAr 1182  
 |||||  
 5665 GGGTGGAAAGCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCG 5714  
 1182 gLeuValProTyrSerAsnAsn 1190  
 |||||  
 5715 CGTGGTTCCTTACTCTGTCATAAC 5739

seq\_name: gb\_pat:AX052638

seq\_documentation\_block:  
 LOCUS AX052638 7873 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 23 from Patent WO0071726.  
 ACCESSION AX052638  
 VERSION AX052638.1 GI:12226928

KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus  
 Viruses; Retrovirdae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.

REFERENCE  
 AUTHORS Federspiel,M.J.  
 TITLE Methods to inhibit infectious agent transmission during  
 xenotransplantation

JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;

Mayo Medical Ventures (US)

FEATURES  
 Location/Qualifiers

1..7873  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 2200 a 1914 g 1798 t  
 ORIGIN

alignment\_scores:  
 Quality: 5862.00 Length: 1194  
 Ratio: 5.106 Gaps: 4  
 Percent Similarity: 96.147 Percent Identity: 94.137

alignment\_block:

US-09-171-553B-5 x AX052638

Align seg 1/1 to: AX052638 from: 1 to: 7873

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
 1637 GGAAGAGCGGGTTCGACCCCTCCCGAGCCCGGTAACCTTGAAGGT 1686

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
 1687 GGAGGGGCAACCAAGTTGAGTCTCGTTGATACCGGAGCAACATTTCAG 1736

34 allLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50  
 1737 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCCTGGGTGATG 1786

51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgArgThrValAs 67  
 1787 GGTGCCACAGGGCAACAACAGTAGTATCCATGGACTTACCCGAAGAACAGTTGA 1836

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
 1837 CTGGAGAGTGGGACGGGTAAACCCACTCGTTCTGGTCATACCTGAGTGCC 1886

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
 1887 CAGCACCCCTCTTAGGTAGAGACTTATTACCAAGATGGGACACAAAT 1936

101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleTh 117  
 |||||

1937 TCTTTTGAACAAGGAAACACAGAGTGTCTGCAAAATAACAAACCTATCAC 1986  
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
 |||||  
 1987 TGTGTTTACCCCTCCAATTAGATGACGAATATGACTATATCTCTCCCTAG 2036  
 134 allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
 |||||  
 2037 TAAAGCCTGATCAAAATATACAAATCTGGTTGGAACAGTTTCCCAAGCC 2086  
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
 |||||  
 2087 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAAGT 2136  
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
 |||||  
 2137 TATTCAACTGAAGGCCAGTGCACACCACTGTCTGATCAGTACAGACAGTACCCCT 2186  
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
 |||||  
 2187 TGAGTAAAGAGCTCAAGAGGAATTCGGCCGCATGTCCAAAGATTAATC 2236  
 201 GlnGlnIleLeuValProValGlnSerProTyrAsnThrProLeuLe 217  
 |||||  
 2237 CAACAGGGCATCCTAGTTCTCTGCAATCTCCCTGGAATACCTCCCTGCT 2286  
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
 |||||  
 2287 ACCGGTTAGAAAGCCCTGGGACTTAATGACTATCGACAGTACAGGACTTGA 2336  
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
 |||||  
 2337 GAGAGTCAATTAACGGGTGCAGGATATACACCAACAGTCCCGAACCCCT 2386  
 251 TyrAsnLeuLeuSerAlaLeuProGluProGluArgAsnTrpTyrThrValle 267  
 |||||  
 2387 TATAACCTCTTGTGTCTCTCCACCCCAACGGAGCTGATATACAGTATT 2436  
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
 |||||  
 2437 GGACTTAAAGGATGCTTTTCTGCTGAGATTACACCCCACTAGCCCAAC 2486  
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
 |||||  
 2487 CACTTTTTCCTTCCAAATGGAGATCCAGGTACGGGAAGAACCCGGGAG 2536  
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
 |||||  
 2537 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 2586  
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
 |||||  
 2587 TGCAAGAGCCCTACACAGAGACTGCGCAACTTCAGAGATCCAACACCCTC 2636  
 334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
 |||||  
 2637 AGGTGACCTCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCAC 2686  
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
 |||||  
 2687 AAACAGGACTGCTTAGAAAGCAGCAAGCACTACTGCTGGAATTGCTGA 2736  
 367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
 |||||  
 2737 CCTAGGCTACAGAGCCTCTGCTAAGAGGCCAGATTTGCAGAGAGAGAG 2786  
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
 |||||  
 2787 TAACATACTTGGGTACAGTTTGGGGACGGGACCGATGCTGACGGAG 2836  
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
 |||||  
 2837 GCACGGAAGAAACCTGTAGTCCAGATACCGGCCCCCAACACACAGCCAAACA 2886

417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
2887 AGTGAGAGAGTTTGGGACAGCTGGATTTCAGACTGTGGATCCCGG 2936  
434 lypPheAlaThrLeuAlaLaProLeuTyrProLeuThrLysGly 450  
|||||  
2937 GGTTCGCGACCTTAGCAGCCCTCTACCCACTAACCAAGAAAGGG 2986  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
2987 GAATTCCTCGGCTCCCTGAGCAGCAGAGGCAATTTGATGCTATCAAAA 3036  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
3037 GGCCTGCTGAGCGCACCTGCTCTGCGCCCTCCCTGACGTAACCTAAACCT 3086  
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500  
3087 TTACCCCTTTATGATGAGCGTAAGGGAGTAGCCCGGGAGTTTAAACC 3136  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
3137 CAAACTCTAGGACCATGGAGGAGACTGTGCTACCTACCTGTCACCAAGACT 3186  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
3187 CGATCCTGTAGCCAGTGGTGGCCCGTATGCCCTGAAGGCTATCGCAGCTG 3236  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
3237 TGCCCTACTGCTCAAGGACGCTGACAAATTGACTTTGGGACAGAAATATA 3286  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
3287 ACTGTAAATAGCCCCCATCGTTGGAGAACATCGTTTCGACGCCCCAGAG 3336  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
3337 CCGATGGATGACCAAGCCCGCATGCCACTATCAAGCCCTGCTCTCA 3386  
584 hGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
3387 CAGAGAGGTTCAGTTCGCTCCACGACGCGCTCTCAACCCCTGCCACTCTT 3436  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
3437 CTGCTGAAGACACTGATGACACCTGACTCATGATTGCCATCACTATT 3486  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
3487 GATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 3536  
634 lylGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
3537 GAGAAGTGTAACTTGTCTACCTGACGGAAGCAGCTATGATGGAAGGT 3586  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
3587 AAGAGATGGCTGGGCGCGGTGGTGACGGGACCGGACCGCATCTGGGC 3636  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
3637 CAGCAGCCTGCCGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGGCC 3686  
684 euThrGlnAlaLeuArgLeuAlaGluClyLysSerIleAsnIleTyrThr 700  
3687 TCACGCAAGCTTGGCGCTGGCGGAAGGAAATCCATAAACATTTATACA 3736  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
3737 GACAGCAGGTATGCCCTTTCGACTGACACACGTACACGGGGCCCATCTATAA 3786  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
3787 GCAAGGGGGTTCGTTACCTCAGCAGGGAGGAAATAAAGAACAAAGAG 3836  
734 luleLeuSerLeuLeuAlaLeuHisLeuProLysArgLeuAlaIle 750  
3837 AAATTTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 3886  
751 lIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLysAs 767  
3887 ATACACTGTCTCGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 3936  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
3937 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCCGAGGGTGTAAAGCTTC 3986  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
3987 TGCCTATGATAGAAACACCCAAAGCCCAAGACCCCGGACGACAGTACAC 4036  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
4037 CTAGACACTGGCAAGAGATAAAAGATAGACCAAGTCTCTGAGACTCC 4086  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
4087 GGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 4136  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
4137 AAGGTTAGATAATGCTCAACAGATACATCTGTACCCACCTAGGAACT 4186  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
4187 AACACCTGCAGCAGTTGGTCAGAAACATCTCTTATCATGTCTTGAGGCT 4236  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 883  
4237 ACCAGGAGTGGCTGATTGGTGGTCAAAACACTGTGTGCCCTGCCAGCTGG 4286  
884 ValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySe 900  
4287 GTA..... 4289  
900 rHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaL 917  
4290 .....AAGCCGCTA 4299  
917 ySTyrGlyAsnLysTyrLeuValPheValAspThrPheSerGlyTrp 933  
4300 AATACGGAACAAATATCTATTTGTTTGTAGACACCTTTTCAGGATGG 4349  
934 ValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLy 950  
4350 GTAGAGCTTATCCTACTAAGAAAGACACTCAACCGTGGTGGCTTAGAA 4399  
950 sIleLeuGlu.GluIlePheProArgPheGlyIleProLysValIleGly 966  
4400 AATACTGGAGGAAATTTTCCAAAGATTTGGAATACCTAAGGTAAATAGG 4449  
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983  
4450 TCAGACAAATGGTCCAGCTTTTCGTTGCCAGGTAAGTACAGACCCCAAGCT 4499  
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS 1000  
4500 GATATTGGGATTGATTGGAAACTGCTTTGTGCATACAGACCCCAAGCT 4549  
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016  
4550 CAGCAGAGGTAGAGAGGATGAATAGAACCACTTAAAGAGAGACCCCTTACATAA 4599  
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033



284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3608 CACTTTTTCCTCGAATGAGAGATCCAGGTACGGGAAGAACCAGCGCAG 3657  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3658 CTCACCTGACCGCGACTCCGCCAAGGGTTCAAGAACTCCCGACCATCTT 3707  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3708 TCACCAAGCCCTACACAGAGACCTGGCCCAACTTCAGGATCCAAACCCCTC 3757  
334 InValThrLeuLeuGlnTyrValAspLeuLeuAlaGlyAlaThr 350  
|||||  
3758 AGGTGACCCCTCCCTCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3807  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuSerAs 367  
|||||  
3808 AAACAGGACTGTTAGAAGGCACGAAGGCACCTACTGCTGGAATGCTCTGA 3857  
367 pleuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
|||||  
3858 CCTAGGCTACAGAGCCTCTGCTAGAAGGCCCGACGATTTGCAGGAGAGAGG 3907  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
|||||  
3908 TAACATACCTGGGTACAGTTTGGGGGACGGCAGCGATGCTGACGGAG 3957  
401 AlaArgLysLysThrValValGlnIleProAlaProThrAlaLysG1 417  
|||||  
3958 GCACGGAAAGAACTGTAGTCCAGATACCGGCCCGCCCAACACAGCCAAACA 4007  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
4008 AGTGAGAGAAATTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 4057  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
4058 GGTTTTGCAGACCTTAGCAGCCCACTCTACCCACTAACCCAAAGAAAAGGG 4107  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
4108 GAATTTCTCTGGGCTCCCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 4157  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
4158 GGCCTGTGACGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCCCT 4207  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
4208 TTACCCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGGAGTTTAAACC 4257  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
4258 CAAACTCTAGGACCATGAGGAGACCTGTGCTACCTACCTGCTCAAGAAGCT 4307  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIav 534  
|||||  
4308 CGATCCTGTAGCCAGTGGTTGGCCCGTATGCTGAAGGCTATCGCAGCTG 4357  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
4358 TGGCCATACCTGCTCAAGGACGCTGACAAATTGACTTTTGGACAGAAATATA 4407  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
4408 ACTGTAATAGCCCCCCTGCTGGAGAACATCGTTTCGACGCCCCCGAGA 4457  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
4458 CCGATGGATGACCAACGCCCGCATACCCACTATCAAGGCCCTGCTCTCA 4507  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
4508 CAGAGAGGGTCACGTTTCGCTCCACACGCGCTCTCAACCTCCGCACTCTT 4557  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
4558 CTGCTTGAAAGAGACTCATGAACACGAGTCACTGATGGCATCAACTATT 4607  
617 urileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4608 GATTGAGGAGACTGGGTCGCCAAGGACCTTACAGACATACCGCTGACTG 4657  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4658 GAGAAGTGTTAACCTGTTTCACTGACGGAAGCAGCTATGTGTGGAAGT 4707  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4708 AAGAGATGGCTGGGCGGGGGTGGTGGACGGGACCCGACGATCTGGGC 4757  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
4758 CAGCAGCCTGCCGGAAGGAACCTTCAGCACAAAAGGCTGAGCTCATGGCCC 4807  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4808 TCACGCAAGCTTTGCGCTGGCCGAAAGGAATCCATAACATTTATACA 4857  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4858 GACAGCAGTATGCTTTGCGACTGCACAGTACACGGGGCCATCTATAA 4907  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4908 GCAAGAGGGGTTGCTTACCTCAGCAGGGGGAATAAAGAACAAAGAGG 4957  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
4958 AAATTTTAAAGCCATTAGAACCTTACATTTGCCAAAAGGCTAGCTATT 5007  
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLysAs 767  
|||||  
5008 ATACACTGCTCTGGACATCAGAAAGCAAGATCCCATATCCAGAGGGAA 5057  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
|||||  
5058 CCAGATGGCTGACCGGTTGCCAAGCAGGCGCCAGGGTGTAAACCTTC 5107  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
5108 TGCTATGATAGAAACACCCAAAGCCCGACAGCCCGGACGACAGTACACC 5157  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5158 CTAGAAGACTGGCAAGAGATAAAAAAGATAGACAGTCTCTCGAGACTCC 5207  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
|||||  
5208 GGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCCTGCCCCCAAG 5257  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||  
5258 AAGGTTTAGAATATGTTCCACAGATACATCGTCTAACCCACCTAGGAAC 5307  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
5308 AAACACCTGCAGCAGTTGGTCAGAACATCTCTTATCATGTTCTGAGGCT 5357  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
|||||  
5358 ACCAGGAGTGGCTGATTGGTGGTGAACAACTGTGTGCCCTGCCAGCTGG 5407  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900

|||||  
5408 TTAATGCTAATCCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAAGC 5457  
901 HisProGlyValAlaHisTipGluValAspPheThrGluValLysProAlaLy 917  
|||||  
5458 CACCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5507  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
5508 ATACGAAACAAATATCTATTGGTTTGTGTAGACACCTTTTCAGGATGGG 5557  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
5558 TAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5607  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
5608 ATACTGGAGGAATTTTCCAGATTTGGAATACCTAAGGTAATAGGGTC 5657  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
5658 AGACAAATGGTCCAGCTTTCGTTGCCAGGTAAGTCAGGGACTGGCCAAAG 5707  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
5708 TATTGGGGATTAAATTGGAAACTGCAATTGTGCATACAGACCCCAAGCTCA 5757  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
5758 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGAGACCCCTTACTAAAT 5807  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
5808 GACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTGTGC 5857  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
5858 TTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
5908 TTACTCTACGGGGACCCCCCAATTGGTAGAAATGCTTCCTCCGTACATAG 5957  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
5958 TGCTGACGTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCACCTTG 6007  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
|||||  
6008 AGTGGGTGACACAACGAGCGTGGAGCAACTCCGGGAGCGCTACTCAGGA 6057  
1100 GlyGlyAspLeuGlnVal 1105  
|||||  
6058 GGAGGAGACTTGCAGATC 6075







|||||  
51 GAAAGACTGAGAAATCCCTTAAAGCTTCGCTCCATCGCTGCTCCTTA 100  
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
|||||  
101 CTCTAAACAATACTCCCGAGCCAGTAGTAAACGCCCTTATAGACAGCTCG 150  
51 AspProHisArgProLeuSerLeuThrTrpLeuIleAspProAspTh 67  
151 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTAGCCCTGATAC 200  
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
|||||  
201 GGGTGTCACTGTAATAGACCTCGAGGTGTCTCTAGAGCACCTGGT 250  
84 rpProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
251 GGCCTGAACCTGCAATTTCTCCCTCGGATTGATTAAACCCGCTGTTAAAGC 300  
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
301 ACACCTCCCAACCTAGTCGCTAGTTATGGGTTCATTGCTGCCCGCAGGCAC 350  
117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgAtgt 134  
351 AGAGAAGAGAATACTGTGGGGTCTCGGGAATCCCTTCTGTAGAGAT 400  
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
401 GGAGCTGCGTCACTCCACGATGGAGACTGGAAATGCCGATCTCTCTC 450  
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIly 167  
451 CAGGACCGGGTAAATTCCTCTTGTTCATATCCGCCCGGCAAGTACAA 500  
167 smMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
501 AATGATGAACCTATATAAGATATAAGACTGCTCCCCATCAGACTTAGATT 550  
184 yrLeuLysIleSerPheThrGluArgLysThrGlyLysTyrSerLysVal 200  
551 ATCTAAAGATAAGTTTCACTGAAAGGAAACAGGAAATATTTCAAAAGTG 600  
201 AspLysTrpTyrGluLeuGlyAsnSerPheLeuLeuTyrGlyGlyAla 217  
601 GATAAATGGTATGAGCTGGGAATAGTTTTTATTATATGCGGGGGAGC 650  
217 aGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluProp 234  
651 AGGGTCCACTTTAACCATTCGCTTAGGATAGAGAGGGGACAGAACCCC 700  
234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAla 250  
701 CTGTGCAATGGAGCCCGATAAGTACTGGCTGAACAGGGGCCCGCCGCC 750  
251 LeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267  
751 CTGGAGCCACCCATAACTTGGCGGTGCCCAATTTAACTCGCTCGCGGCC 800  
267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284  
801 TGACATAACACAGCCGCTAGCACACATGATCCACTGGATTGATTCCTACCA 850  
284 snThrProArgAsnSerProGlyValProValLysThrGlyClnArgLeu 300  
851 ACACGCCTAGAAACTCCCGAGGTGTCTGTTAAGACAGGACAGAGACTC 900  
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317  
901 TTCAGTCTCATCAGGGAGCTTTTCAAGCCATFCAACTCCACCGACCTGA 950  
317 pAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrTyrG 334  
|||||

951 TGGCACCTTCTTCTTGTGGCTTTGTCTATCTCTCAGGGCCTCCTTATTATG 1000  
334 luGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsn 350  
1001 AGGGGATGGCTAAAGAAAGAAATTTCAATGTGACCAAGAGCATAGAAAT 1050  
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGl 367  
1051 CAATGTACATGGGGTCCGAAATTAAGCTTACCCCTCAGCTGAAGTTTCCGG 1100  
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCyst 384  
1101 GAAGGGGACATGATAGAAAGAGCTCCCCATCCACCAACACACCTTTGCT 1150  
384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400  
1151 ATAGTACTGTGTTTATGAGCAGGCCCTCAGAAATCAGTATTTAGTACCT 1200  
401 GlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValse 417  
1201 GGTATTAACAGGTGGTGGCATGCAATACTGGGTAAACCCCTCTGTGTTTC 1250  
417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434  
1251 CACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCTATGTCCTCAAATCG 1300  
434 alProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAsp 450  
1301 TCCCCCGAGTGTACTACCATCTCAGGAAGTGGTCTTGTATGAATATGAC 1350  
451 TyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVa 467  
1351 TATCGGTATTAACCCAGCAAAAGAGAACCCGCTATCCCTTACCCCTAGCTGT 1400  
467 lMetLeuGlyLeuGlyThrAlaValGlyValIglyThrGlyThrAlaAlaL 484  
1401 AATGCTCGGATTAGGAGCGGCGGTGGCTAGGAACAGGACAGCTGCC 1450  
484 euIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAla 500  
1451 TGATCACAGGACCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCG 1500  
501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuGl 517  
1501 GCCATGACAGAAAGATCTCCGAGCCTTAAAGAGAGTCTGTTAGCAACCTAGA 1550  
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534  
1551 AGAGTCCCTGACTTCTTTGTCTGAAGTGGTCTACAGAACCCGGAGGGGAT 1600  
534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550  
1601 TAGATCTGCTGTTTCTTAAGAGAGGTGGTATTGTGTCAGCCTTAAAAAGAA 1650  
551 GluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAs 567  
1651 GAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGACACTCCATGAA 1700  
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584  
1701 CAAGCTTAGAAAAAAGTTAGAGAGCGCTCGAAGGAAAGAGAGCGTGACC 1750  
584 lnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLeu 600  
1751 AGGGGTGGTTTGAAGGATGGTTCAACAGGTCTCCTTGGATGACACCCCTG 1800  
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThrVa 617  
1801 CTTTCTCTCTGAGGGGGCCCTTAGTAGTCTCTCTCTCTCTCTCTCTCTCT 1850  
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgVals 634  
1851 TGGGCCCTTGCTTAATTAATAGGTTTGTGTGCTTTGTTAGAGACAGAGTGA 1900

```
634 erAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuSer 650
|||||
1901 GTGAGTCCAGATGCTAGTACCTAGGCAACAGTACCAAGGCTCTCTGAGC 1950

651 GlnGlyGluThrAspLeu 656
|||||
1951 CAAGAGAACTGACCTC 1968

seq_name: gb_pat:A66553
seq_documentation_block:
LOCUS A66553 8209 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 3 from Patent W09740167.
ACCESSION A66553
VERSION A66553.1 GI:4538106
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
Virus; Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 8209)
AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 3 30-OCT-1997;
Q ONE BIOTECH LTD (GB)
FEATURES
Location/Qualifiers
1..8209
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
588..2162
/codon_start=1
/product="VIRAL CORE PROTEIN"
/protein_id="CAB39347.1"
/db_xref="GI:4538107"
/translacion="MGQTVTPPLSLTDHWTETVRSRAHNSLVQVKKGPQWTFCASEWP
TFDQWSEGTNSBIIILAVKAIIFGTGSHPDQEPYILTWDLAEDPPPPWPKWFLN
KPKFGPRILALGKKNHSAEKVESSYLPDRGAADLAGTPTCPTPLSGCCGEG
TSAPGAPVEGPAAGTTSRRGATPRTDEIAILPLRTYGPMPGQLOLPYWPFS
ADLYNKNTHNPFSEDRDLTGLVESLMFHOPTWDDCOILLOTFTTEERILLEA
KNVFCAGDRPTOLONEIDMFPLTPGWDYNTAEGRESLKIYRQALVAGLARSRR
TNLVRVMQGPNEPPSVFLERLMEAFRRFTFPDPTSEAQKASVALAFIQSALDIR
KLOLEGLQEAELDLVREAEKVVYRETEEEKEQREKEEREERDRRQENLT
KILAAVVEGKSSRERERFRKIRSGPRQSGNLNRTPLDKQOCAYCKEKGHWARNCPK
KGNKGPVLALEEDKD"
<2163..5747
/codon_start=1
/product="REVERSE TRANSCRIPTASE"
/protein_id="CAB39348.1"
/db_xref="GI:4538108"
/translacion="GRRGSDPLPEPRVTLKVEGPVEFLVDTGAEHSLVLOPLCKLKE
KSWMGATGQRYPTWTRRTVYDVGVRTHSFLVIPCEPVLGLRDLTLTKNGQISF
EGREVSNNKAPITVLIQLDDETRLSYQVKPDQDQISWLEQPPQAWAETAGMGLA
KQVPQVTLKASATPSVROQDPSREAREGIMPHVQLIOQILVPOSPWNTPLLP
VRKPGTNDYRPVQDRIYVKNRQDIHTPNPNNLLSALPPERNWYTLVDLKDAPFLC
RLHPTSQLFAPRPDQRTGQTLTIRLPQGFKNSTIFDEALHRLANFRIOHPO
VTLQVDDLLLAGATKODCLEGTKALLLSLDGYSASAKKQICREVTYLYSLR
GQORWLTARKTVQVQIPATTAQVRELGTCAGCRWIPGFIATLAPLPLPKXC
GFSWAPHQKADAIKALLSAPALAPDVTKPFYLYDERKGVARGVLTQTLPWRR
PVALYSLKLDPVASGWPVCLKAIAVAAILVKDADKLTLQNLITYAHALLENIVRPP
KDLTDIPITGEVLTWFTDGSVVEGKRMAGAAVVDGRTIWASSLPGTSQAQALM
ALQALRLAEGKSINITYDSYAFATAHVGAIIYKORGLLT SAGREIKNKEIISLE
ALHLPKRLAIHCPGHOKADLIISGQNMADRVAKQAQAVNLLPIETPKAPPRQ
YILEWQEIKKIDQFSETPEGTCTYSKREILPHKEGLEIYQQIHRUHLTGLKHLQ
VRTSPYHLRIPGVADSVKHCPCQLNANPPIPPGKRLRSGHPGAHWEDVTEV
PAKYNGKYLKLVFDVTSQWAEYPTKSTVAVKILIEEIPFRGIPKVGSDNGPA
FVAQVSGQAKLTGIDMKLHCAVPOSSGOVERNMTIKETLTKLTETGNDMALL
PPVLFVRNTPGQFGLTPYELLYGGPPLETAFAHSADVLLSLOPLSRILAEWVQ
RAWKQLREAYSGDQLQVPHRFQVGSVYVRHRAGNLETRKNGKLYVLLLTPTAVKVE
GLPTWISHAVKVPAPPDPSGWRAEKTEPNLRLHRLVFPYSNNNSPGQ"
5620..7590
/codon_start=1
```

```
/product="VIRAL ENVELOPE PROTEIN"
/protein_id="CAB39349.1"
/db_xref="GI:4538109"
/translacion="MHPTLSRRHLPTRGGEPRKLRIPLSFASIAWFLTLTITPOASSK
RLIDSSNPHRPLSLTLWLIIDPDGTVTYNSTRGVAPRTWHPPELHFLRLINPAVKSTP
PNLVSFGFYCCPGTEREKYCGGSEFCRWSCVTSNDGDKWPIQLQDRVAFSFPN
SPGKYKMKLYDKSCPSDLXYLSFTEKTKGYSKVDKWEYELGNSFLLYGGAG
STGLIRLRIETGTPPVAMPDKVLAEQGPALPEPHNLVPQLTSLRPDITOPPSNS
TTGLIPNTPRNSPCVPKTKORFLSLIOGAFOAINSTDPTDSCWCLSSGPPYVE
GMAKERFNTVKEHNOCTWGSRNKLLTTEVSGKGTGIGRAPPSHOHLCYSTVVYEOA
SENQYLPVGYNRWACNGLTPCYSTSVFNQSKDFCVWVQIVPRVYHPREVILDEVD
YRNRKPREPSLTALVMLGTVAGVGTAAALITGPQQLKEKGLGEHAAETDLRA
LKESVNLNEESLTSLSEVYLNQRRGLDLFLREGGLCAALKECCFYVDHSGAIRDSM
NKLKKLLERREREADQGWFEGRNRPWMTTLLSALTGPLVVLVLLLLVVGCLINR
FVAFVRRVSAVQIMVLVROQVGLLSOGETDL"
BASE COUNT 2168 a 2064 c 2154 g 1823 t
ORIGIN
alignment_scores:
Quality: 3515.00 Length: 656
Ratio: 5.358 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-171-553B-6 x A66553
Align seg 1/1 to: A66553 from: 1 to: 8209
1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGlyGluPr 17
|||||
5620 ATGCATCCACAGTTAAGCGGGGCCACCTCCGACTCGGGGTGAGAGCC 5669
17 OlyArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
5670 GAAAGACTCAGAATCCCTTAAGCTTCGCCTCCATCGCTGTTCTCTTA 5719
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSer 50
|||||
5720 CTCACATAACTCCCGAGCCAGTAGTAAGCCCTATATACAGCTCG 5769
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleAspProAsp 67
|||||
5770 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 5819
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
5820 GGGTGTCACTGTAATAATAGCACTCGAGGTGTTGCTCTAGAGCACCTGT 5869
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5870 GGCTGAACCTGCAATTCTCCCTCCGATGATTAACCCCTGTAAAAAGC 5919
101 ThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysProGly 117
|||||
5920 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAGCAC 5969
117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArg 134
|||||
5970 AGAGAAGAGAATACTGCGGGGTTCGGGGAATCCCTTCTCTAGAGAT 6019
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
6020 GGAGCTGCTGCACCTCCACAGATGAGACTGGAAATGGCCGATCTCTC 6069
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
|||||
6070 CAGGACCGGTAAAAATTCCTTTTGTCAATTCCGCGCCGGGCAAGTACAA 6119
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
|||||
6120 AATGATGAACATATATAAGAAAGAGCTGCTCCCATCAGACTTAGATT 6169
```

184 yrluLysIleSerPheThrGluArgLysThrGlyLysTyrSerLysVal 200  
6170 ATCTAAGATAAGTTTCACTGAAGGAAACAGAGAAATATTTCAAAAGTG 6219  
201 AspLysTrpTyrGluLeuGlyAsnSerPheLeuLeuTyrGlyGlyAl 217  
6220 GATAAATGATGACCTGGGGAATAGTTTTTATTATATGCGGGGGAGC 6269  
217 aGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluPro 234  
6270 AGGGTCCACTTTAAACCATTCGCCTTAGGATAGAGACGGGACAGAACCC 6319  
234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProAla 250  
6320 CTGTGGCAATGGGACCCGATAAGTACTGGCTGAACAGGGGCCGCCGCC 6369  
251 LeuGluProPheHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267  
6370 CTGGAGCCACCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGGGCC 6419  
267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284  
6420 TGACATAACACAGCGCCCTAGCAACAGTACCACCTGGATTGATTCCTACCA 6469  
284 snThrProArgAsnSerProGlyValProValLysThrGlyGlnArgLeu 300  
6470 ACACGGCTAGAACATCCCGAGGTGTCTCTGTTAAGACAGGACAGACTC 6519  
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317  
6520 TTCAGTCTCATCCAGGAGCTTTCCAGGCCATCACTCCACCGACCCCTGA 6569  
317 pAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrTyrG 334  
6570 TGCCACTCTCTCTGTGGCTTTGTCTATCTCTCAGGGCTCTCTATTATG 6619  
334 luGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArgAsn 350  
6620 AGGGATGGCTAAGAAAGAAATTCATGTGACCAAGAGAGCATGAAT 6669  
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSerG 367  
6670 CAATGTACATGGGGTCCCGAAATAAGCTTACCTCACTGAAGTTTCCGG 6719  
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCyst 384  
6720 GAAGGGAGCATCATAGGAAAGCTCCGCCATCCCAACCAACACCTTTGCT 6769  
384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400  
6770 ATAGTACTGTGGTTATGAGCAGGGCTCAGAAAATCAGTATTAGTACCT 6819  
401 GlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSe 417  
6820 GGTATACAGGTTGGTGGCATGCAATAGCTGGTTAACCCCTGTGTTTC 6869  
417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434  
6870 CACCTCAGTCTTCAACCAATCCAAAGATTTCTGTCTATGTCGTCCTG 6919  
434 aLProArgValTyrTyrHisProGluGluValValLeuAspGluThrAsp 450  
6920 TCCCCGAGTGACTACCATCTCTGAGGAAGTGGTCTGTATGAATATGAC 6969  
451 TyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaV 467  
6970 TATCGGTATAACCGACCAAAAGAGAACCCGTATCCCTTACCTAGCTGT 7019  
467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaL 484  
7020 AATGTCGGATTAGGACCGCGCTGGCTAGGAACAGGACAGCTGCC 7069  
484 euIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGlnLeuHisAla 500

7070 TGATCACAGGACCACACAGCTAGAGAAAGAGCTTGGTGAGCTACATCG 7119  
501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuG 517  
7120 GCCATGACAGAAAGATCTCCGAGCCTTAAGAGAGTCTGTTAGCAACCTAG 7169  
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534  
7170 AGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAGAACCGGAGGGAT 7219  
534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550  
7220 TAGATCTGCTCTTCTTAAGAGAGGTGGTATGTGCAGCCTTAAGAAGAA 7269  
551 GluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAs 567  
7270 GAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAA 7319  
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584  
7320 CAAGCTTAGAAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGGCTGACC 7369  
584 lGlyTrpPheGluGlyTrpPheAsnArgSerProTyrMetThrThrLeu 600  
7370 AGGGTGGTTGAAGGATGGTTCAACAGGCTCTCTGGATGACCAACCTG 7419  
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThrVa 617  
7420 CTTTCTCTCTGACGGGCGCCCTAGTAGTCTGCTCTCTGTTACTTACAGT 7469  
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValS 634  
7470 TGGCCCTTGCTTAAATAGTGTGTGGCTTTGTTAGAGAACGAGTGA 7519  
634 erAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSer 650  
7520 GTGAGTCCAGATCATGTTACTTAGGCAACAGTACCAGGCGCTTCTGAGC 7569  
651 GlnGlyGluThrAspLeu 656  
7570 CAAGGAGAACTGACCTC 7587  
seq\_name: gb\_pat:A66552  
seq\_documentation\_block: 8196 bp DNA PAT 29-MAR-1999  
LOCUS A66552  
DEFINITION Sequence 2 from Patent WO9740167.  
ACCESSION A66552  
VERSION A66552.1 GI:4538105  
KEYWORDS  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
REFERENCE 1 (bases 1 to 8196)  
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.  
TITLE PORCINE RETROVIRUS  
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;  
Q ONE BIOTECH LTD (GB)  
FEATURES  
source  
1..8196  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"  
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others  
ORIGIN

alignment\_scores:  
Quality: 3509.00 Length: 656  
Ratio: 5.357 Gaps: 0  
Percent Similarity: 99.848 Percent Identity: 99.848

## alignment\_block:

US-09-171-553b-6 x A66552

Align seg 1/1 to: A66552 from: 1 to: 8196

1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGlyGluPr 17  
 5606 ATGCATCCACGTTAAGCYGGCCACCTCCCGACTCGGGTGGAGACC 5655  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 5656 GAAAAGACTGAGATCCCTTAAAGCTTCGCCTCCATCGCTGGTTCCTTA 5705  
 34 hrLeuThrIleThrProGlnAlaSerLysArgLeuIleAspSerSer 50  
 5706 CTCTAACAAATAACTCCCGAGCCAGTAGTAAACGCCATTATAGACAGCTCG 5755  
 51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67  
 5756 AACCCCCATAGACCTTTATCCCTTACCTGGCTGATTAATTGACCTGATAC 5805  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 5806 GGGTGTCACTGTAATAGCATCTCAGAGGTGCTCCTAGAGGCACTCGGT 5855  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 5856 GGCCTGAACCTCAATTCCTGCCCTCCGATGATTAAACCCGCTGTTAARAGC 5905  
 101 ThrProAsnLeuValArgSerTyrrGlyPheTyrrCysCysProGlyTh 117  
 5906 ACACCTCCCAACCTAGTCCGTAGTATGGGTCTTATTGCTGCCAGGCAC 5955  
 117 rGluLysGluLysTyrrCysGlyGlySerGlyGluSerPheCysArgArgT 134  
 5956 AGAGNAAGAGAAATACTGTGGGGTTCGGGGAAATCCTTCGTAGGAGAT 6005  
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 6006 GGAGCTGGCTCACCTCCCAACGATGGAGACTGGAATGGCCGATCTCTCTC 6055  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrrLy 167  
 6056 CAGGACCGGGTAAATAATTCCTCTGTCAATTCGGCGCCGGGCAAGTACAA 6105  
 167 sMetMetLysLeuTyrrLysAspLysSerCysSerProSerAspLeuAspT 184  
 6106 AATGATGAACTATATAAGATAAGAGCTGCTCCCCATCAGACTTAGATT 6155  
 184 yrLeuLysIleSerPheThrGluArgLysThrGlyLysTyrrSerLysVal 200  
 6156 ATCTAAAGATAAGTTTCACTGAAAGGAAACAGCAAAATAATTCAAAAGTG 6205  
 201 AspLysTrpTyrrGluLeuGlyAsnSerPheLeuLeuTyrrGlyGlyAl 217  
 6206 GATAAATGATGATGAGCTGGGAAATAGTTTTTATATATGCGGGGGAGC 6255  
 217 aGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluProp 234  
 6256 AGGGTCCACTTTAACCATTCGCCTTAGGATAGACGGGGACAGAACCCC 6305  
 234 roValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAla 250  
 6306 CTGTGGCAATGGGACCGGATAAAGTACTGGCTGAACAGGGGCCCGCCGCC 6355  
 251 LeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPr 267  
 6356 CTGGAGCCACCGCAATACTTCCCGGTGCCCCAATAACCTCGCTGCGGCC 6405  
 267 oAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrA 284  
 6406 TGACATTAACAGCGCGCTACCAACAGTACCACCTGGATTGATTCCTACCA 6455

284 snThrProArgAsnSerProGlyValProValLysThrGlyGlnArgLeu 300  
 6456 ACACGCTAGAACTCCCGAGGTTCCTGTTAAGACAGGACAGAGACTC 6505  
 301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317  
 6506 TTCAGTCTCATCCAGGGAGCTTTCGAAGCCATCAACTCCACCGACCTGA 6555  
 317 pAlaThrSerSerCysTrpLeuSerLeuSerSerGlyProProTyrrTrg 334  
 6556 TGGCACCTTCTCTGTGTGGCTTTGTCTATCTCAGGGCCCTCCTATTATG 6605  
 334 luGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsn 350  
 6606 AGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAGAAAT 6655  
 351 GlnCysThrTrpGlySerArgAsnLysLeuThrIleuThrGluValSerGl 367  
 6656 CAATGTACATGGGGTCCCGAAATAAGCTTACCTCCTCACTGAAGTTCCGG 6705  
 367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCyst 384  
 6706 GAAGGGACATGCATAGGAAAGCTCCCCATCCCAACACCTTTGCT 6755  
 384 yrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400  
 6756 ATAGTACTGTGTTTATGACAGCGCTCAGAAATCAGTATTATTAGTACCT 6805  
 401 GlyTyrrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValse 417  
 6806 GGTATATAACAGGTGGTGGGCATCACTACTGGGTAAACCCCTGTCTTC 6855  
 417 rThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnIleV 434  
 6856 CACCTCAGTCTTCAACCAATCCAAAGATTTCGTGTCATGGTCCAAATCG 6905  
 434 alProArgValTyrTyrrHisProGluGluValValLeuAspGluTyrrAsp 450  
 6906 TCCCCCGAGTGTACTACCATCCTGAGGAAGTGTCTCTGATGAATATGAC 6955  
 451 TyrrArgTyrrAsnArgProLysArgGluProValSerLeuThrLeuAlaVa 467  
 6956 TATCGGTATAAACGACGACCAAGAGAACCGGTATCCCTTACCTAGCTGT 7005  
 467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaL 484  
 7006 AATGCTCGGATTAGGGACGCGCTGGCGTAGGAACAGGGAGCTGCC 7055  
 484 euIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAla 500  
 7056 TGATCAGAGCACACAGCAGCTAGACAAAGGACTTGGTGAGCTACATGCG 7105  
 501 AlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuGl 517  
 7106 GCCATGACAGAGATCTCCGAGCTTAAAGGAGTCTGTTAGCAACCTAGA 7155  
 517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyL 534  
 7156 AGAGTCCCTGACCTCTTGTCTGAAGTGGTTCACAGAACCGGAGGGAT 7205  
 534 euAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGlu 550  
 7206 TAGATCTGCTGTTTCTAAGAGAGGTGGGTATTGTGCAGCCTTAAAGAA 7255  
 551 GluCysCysPheTyrrValAspHisSerGlyAlaIleArgAspSerMetAs 567  
 7256 GAATGTTGCTTCTATGTAGTACATCAGGAGCCATCAGAGACTCCATGAA 7305  
 567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584  
 7306 CAAGCTTAAAAAAGTTAGAGAGGCGCTCAAGGGAAAGAGAGGCTGACC 7355  
 584 lnglyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLeu 600

|||||  
7356 AGGGGTGGTTTGAAGGATGGTTCAACAGGTCCTCGTGGATGACCAACCCCTG 7405  
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThVa 617  
7406 CTTTCCTGCTCTGACGGGGCCCTAGTAGTCCCTGCTCTTACTTACAGT 7455  
617 lGlyProCysLeuLeuAsnArgPheValAlaPheValArgGluArgVals 634  
7456 TGGGCTTCTCTTAATTAATAGCTTGTGCTTTGTAGAGAACAGTGA 7505  
634 exAlaValGlnLeuMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSer 650  
7506 GTGCAGTCAGATCATGTGTTAGTACACAGTACACAGGCTTCTGAGC 7555  
651 GlnGlyGluThrAspLeu 656  
7556 CAAGGAGAACTGACCTC 7573

seq\_name: gb\_pat: A66559

seq\_documentation\_block: 1974 bp DNA PAT 29-MAR-1999

LOCUS A66559 Sequence 9 from Patent WO9740167.

ACCESSION A66559

VERSION A66559.1 GI:4538112

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1974)

AUTHORS Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

TITLE PORCINE RETROVIRUS

JOURNAL Patent: WO 9740167-A 9 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES Location/Qualifiers

source 1..1974

/organism="unidentified"

/db\_xref="taxon:32644"

BASE COUNT 534 a 489 c 483 g 468 t

ORIGIN

alignment\_scores:

Quality: 3455.00 Length: 658

Ratio: 5.291 Gaps: 2

Percent Similarity: 99.240 Percent Identity: 99.088

alignment\_block:

US-09-171-553B-6 x A66559 ..

Align seg 1/1 to: A66559 from: 1 to: 1974

1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGluPr 17

1 ATCATCCCCAGGTTAAGCTGGGGCCACCTCCCGACTCGGGGTGGAGAGCC 50

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34

51 GAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTA 100

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuLeuAspSerSer 50

101 CTCTAAACAATACTCCCGAGGCGAGTAGTAAACGGCTTATAGACAGCTCG 150

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleLeuAspProAspTh 67

151 AACCCCATAGACCTTTATCCCTTACCTGCTGATTTATGACCTGATAC 200

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84

201 GGGTGTCACTGTAATAGCACCTGAGGTGTGTCTCTAGAGGACACCTGCT 250

84 rpProGluLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100  
251 GGCCTGAACGCAATTTCTGCTCGATTGATTAAACCCGCTGTAAAGC 300  
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
301 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTTGCTGCCAGGCAC 350  
117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT 134  
351 AGAGAAAGAGAAATACTGTGGGGTTCTGGGAATCCTTCTGTAGGAGAT 400  
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProLleSerLeu 150  
401 GGAGCTCGTCCACCTCCACAGATGGAGACTGGAAATGGCCGATCTCTCTC 450  
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
451 CAGGACCGGGTAAATTTCTCTTTGTCAATTCGGCCCCGGGCAAGTACAA 500  
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
501 AATGATGAAACTATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 550  
184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200  
551 ATCTAAAGATAAGTTTCACTGAAAGGAAACAGAGAAATATTCAAAAG 600  
200 aAspLysTyrTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216  
601 TGGATAAATGGTAAGAGCTGGGGAATAGTTTTTTTATAAATATGGCGGGG 650  
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233  
651 AGCAGGTCCACTTAACTTCCCTTAGGATAGAGACGGGACAGAAC 700  
233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249  
701 CCCCTGTGGCAATGGAGCCGATAAAGTACTGGCTGAACAGGGGCCCCCG 750  
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266  
751 GCCCTGGAGCCACCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGG 800  
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283  
801 GCCTGACATAACACAGCGCGCTAGCAACAGTACCACCTGGATTGATTCCTA 850  
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299  
851 CCAACAGCGCTAGAAACTCCCGAGGTGTCTCTTAAAGACAGGACAGAGA 900  
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316  
901 CTCCTTCAGTCTCATCCAGGGAGCTTTCCAAGCCATCAACTCCACCGACC 950  
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProTrpTrt 333  
951 TGATGCCACTTCTCTTGTGGCTTTGTCTATCTCTCAGGGCCCTCTTATT 1000  
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349  
1001 ATGAGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGACATAGA 1050  
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366  
1051 AATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTC 1100  
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 383  
1101 CGGGAAGGGGACATGCATAGGAAAGCTCCCGCATCCCAACACACCTTT 1150  
383 yTyrSerThrValValThrGluGlnAlaSerGluAsnGlnTyrLeuVal 399

```
|||||
1151 GCTATAGTACTGCTGTTTATGAGCAGGCTCAGAAATCACTATTACTA 1200
400 ProGlyTyrAsnArgTyrTrpAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
1201 CCTGGTTATACAGGTGGTGGCATGCAATACTGGGTTAAACCCCTGTGT 1250
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
1251 TTCCACCTCAGTCTTCAACCAATCCAAGATTGTGTCTCATGTGCCAAA 1300
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
|||||
1301 TCGTCCCCGAGTGCTACTTACCATCTGAGGAAGTGGCTTGATGAATAT 1350
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
1351 GACTATCGGTATACCGACCAAAAGAGACCCGTATCCCTTACCCCTAGC 1400
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 483
|||||
1401 TGTAATGCTCGATTAGGACGCGCGTGGCGTAGAACACGGGACAGCTG 1450
483 laLeuIleThrGlyProGlnLeuGluLysGlyLeuGlyGluLeuHis 499
|||||
1451 CCCTGATCACAGGACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 1500
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
|||||
1501 CGGCCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTCTGTTACCACT 1550
516 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533
|||||
1551 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTTCTACAGACCGGAGGG 1600
533 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLys 549
|||||
1601 GATTAGATCTGCTCTTCTAAGAGAAGGTGGGTTATGTGAGCCTTAAAA 1650
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgaspSerMe 566
|||||
1651 GAAGAATGTGCTTCTATGATAGTACTCAGGAGCCATCAGAGACTCCAT 1700
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluGluAlaA 583
|||||
1701 GAACAAGCTTAGAAAAAAGTTAGAGAGCGCTCGAAGGGAAGAGAGCGTG 1750
583 spGlnGlyTyrPheGluGlyTrpPheAsnArgSerProTyrMetThrThr 599
|||||
1751 ACCAGGGGTGGTTTGAAGGATGGTTCACAGGCTCCTTGGATGACCACC 1800
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616
|||||
1801 CTGCTTCTCTGCTGACGGGCCCTAGTAGTCTGCTCTCTGTACTTAC 1850
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
|||||
1851 AGTTGGCCCTTGCTTAATTAATAGTTTGTGCTTGTAGAGAACGAG 1900
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
|||||
1901 TGAGTGAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCCTTCTG 1950
650 SerGlnGlyGluThrAspLeu 656
|||||
1951 AGCAAGAGAAACTGACCTC 1971
```

seq\_name: gb\_pat:AX002804

seq\_documentation\_block:

LOCUS AX002804

DEFINITION Sequence 3 from Patent WO9853104.

ACCESSION AX002804

PAT

21-AUG-2000

VERSION AX002804.1 GI:9885132  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

porcine endogenous retrovirus.  
porcine endogenous retrovirus  
Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.  
1 (bases 1 to 3482)  
Stoye,J.P. and Weiss,R.A.  
Detection of retroviral subtypes based upon envelope specific  
sequences  
Patent: WO 9853104-A 3 26-NOV-1998;  
MEDICAL RES COUNCIL (GB); STOVE JONATHAN PAUL (GB)  
Location/Qualifiers  
1..3482  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

927 a 854 c 867 g 834 t

alignment\_scores:

Quality: 3433.00 Length: 658  
Ratio: 5.257 Gaps: 2  
Percent Similarity: 99.240 Percent Identity: 98.176

alignment\_block:

US-09-171-553B-6 x AX002804 ..

Align seg 1/1 to: AX002804 from: 1 to: 3482

1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17  
911 ATCATCCCCACGTTAGCTGGGCCACCTCCCGACTCGGGGGGAGAGCC 960  
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
961 GAAAGACTCAGAAATCCCTTAAGCTTCGCTCCATCGCTGGTCTCTTA 1010  
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleaspSerSer 50  
1011 CTCTAACAATACTCCCGAGGCCAGTAGTAAACGCGCTTATACAGACTCG 1060  
51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67  
1061 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 1110  
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
1111 GGGTGTCACTGTAATAGACTCGAGGTGTGCTCTAGAGCAGCCTGTGT 1160  
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
1161 GGCCTGAATGCTATTTCTCCCTCCGATTGATTAAACCCGCTGTTAAAGC 1210  
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
1211 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAGGCAC 1260  
117 rGluLysGluLysTyrCysGlySerGlySerGlyGluSerPheCysArgArgT 134  
1261 AGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCCTTCTTAGAGAT 1310  
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
1311 GGAGCTGCTCACCCTCAACAGATGGAGACTGGAAATGGCCGATCTCTCTC 1360  
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
1361 CAGGACCGGGTAAATTTCTCTTTTGTCAATTCGCGCCGCGGCACTACAA 1410  
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
1411 AGTGATGAACCTATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 1460



184 yfLeuLysIleSerPheThrGlu...ArgLysThrGlyLysTyrSerLysV 200  
1461 ATCTAAGATAGTTTCTACTGAAAGGAGAAACAGAGAAATATCAAAAG 1510  
200 aAlAspLysTrpTyrGluLeuGlyAsnSerPheLeu...LeuTyrGlyGlyG 216  
1511 TGCATAAAATGGTATGAGCTGGGGAATAGTTTTTTATATAAATATGGCGGG 1560  
216 vAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233  
1561 ACAGGGTCCACTTAAACCATTCGGCTTAGGATAGAGACGGGACAGAAC 1610  
233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249  
1611 CCCCTGTGGCAGTGGGACCCGATAGTACTGGCTGACAGGGGCCCGG 1660  
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266  
1661 GCCCTGGAGCCACCATTAACCTTGGCGGTGGCCCAATTAACCTCGCTGCG 1710  
266 gProAspIleThrGlnProProSerAspSerThrThrGlyLeuIleProT 283  
1711 GCCTGACATTAACACAGCCGCTAGACCGTACCACCTGGAITGATTCCTA 1760  
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299  
1761 CCACACGCTAGAAATCCCCAGGTTCCTGTTAAGACAGGACAGAGA 1810  
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316  
1811 CTCTTCAGTCTCATCCAGGAGCTTTCGAAGCCATCAACTCCACGACCC 1860  
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333  
1861 TGATGCCACTTCTCTTGTGGCTTGTCTATCCTCAGGGCCCTCCTTATT 1910  
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349  
1911 ATGAGGGGATGCTAAAGAAAGGAAATTCATGTGACCAAGAGCATAGA 1960  
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSe 366  
1961 AATCAATGTACATGGGGTCCCGAAATAGCTTACCCTCAGTGAAGTTTC 2010  
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 383  
2011 CGGGAAGGGACATGCATAGGAAAGCTCCCCCATCCCAACACACCTTT 2060  
383 ysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399  
2061 GCTATAGTACTGTGGTTTATGAGCAGGCTCAGAAAATCAGTATTAGTA 2110  
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416  
2111 CCTGGTTATAACAGGTGGGGCATGCAATCTGGGTAAACCCCTGTGT 2160  
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433  
2161 TTCCACCTCAGTCTCAACCAATCCAAAGATTTCTGTGTCATGGTCCAAA 2210  
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449  
2211 TCGTCCCGGAGTGTACTACCATCTCTGAGGAAGTGGTCTTCGATGAATAT 2260  
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466  
2261 GACTATCGGTATACCCACCAAAAGAGAACCCGTATCCCTTACCCTAGC 2310  
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAla 483  
2311 TGTAATGCTCGGATTAGGACGCGCTTGGCGGTAGGAACAGGAGACAGCTG 2360

483 laLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyGluLeuHis 499  
2361 CCCTGATCAGAGCACACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 2410  
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516  
2411 GCGCCATGACAGAAGATCTCCGAGCCTTAGAGAGTCTGTAGCAACCT 2460  
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 533  
2461 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTACAGAACCGGAGGG 2510  
533 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLys 549  
2511 GATTAGATCTGCTCTTCTTAAGAGAAGTGGGTATGTGCAGCCTTAAAA 2560  
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566  
2561 GAGAATGTTGCTTCTATGATGATCACTCAGAGCCATCAGAGACTCCAT 2610  
566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAla 583  
2611 GAGCAACCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGCTG 2660  
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599  
2661 ACCAGGGTGGTGTGAAGAGTGTTCACAGGTCCTCTGGATGACACC 2710  
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616  
2711 CTGCTTCTCTGACGAGGACCCCTAGTAGTCTCTGCTCTGTACTTAC 2760  
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633  
2761 AGTTGGGCTTGCCTTAATTAATAGTGTGTGCTTGTAGAGAACGAG 2810  
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649  
2811 TGAGTGAGTCCAGATCATGTACTTAGGCAACAGTACCAGGCGCTCTG 2860  
650 SerGlnGlyGluThrAspLeu 656  
2861 AGCCAAGGAGAACTGACCTC 2881  
seq\_name: gb\_vi:PERENV2  
seq\_documentation\_block: 3482 bp RNA VRL 14-OCT-1997  
LOCUS PERENV2 Porcine endogenous retrovirus env gene, 3482 bp.  
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.  
ACCESSION Y12239  
VERSION Y12239.1 GI:2576324  
KEYWORDS env gene.  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus.  
REFERENCE 1 (bases 1 to 3482)  
AUTHORS Le Tissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.  
TITLE Two sets of human-tropic pig retrovirus  
JOURNAL Nature 389 (6652), 681-682 (1997)  
MEDLINE 97478526  
REFERENCE 2 (bases 1 to 3482)  
AUTHORS Stoye,J.P.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical  
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK  
REMARK revised by author 16-SEP-97  
FEATURES Location/Qualifiers  
source 1..3482  
/organism="porcine endogenous retrovirus"  
/specific\_host="Sus scrofa"  
/db\_xref="taxon:61673"  
/note="PERV-B"

gene

911.. .2884

/gene="env"

CDS

911.. .2884

/gene="env"

/codon\_start=1

/protein\_id="CAA72928.1"

/db\_xref="GI:2576325"

/db\_xref="SPTREMBL:O41173"

translation="MHPTLSWRHLPTRGGEKRLRPLSFASIAWFLTLTITPOASSK  
RLDSSNHRPLSLTWLIIDPDGTVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTP  
PNLVRSGYCCPGTEKEKYCGSGESFCRRWSCVTSNDGDKWPLSLQDRVKFSEVN  
SGKTLRLRIETGEPVAVGDDPKVLAEGPPALEPPLHNPVQLTSLRDPDITQPEN  
GTGLIPTNPRNPGVPVKTGQRLFSLIQGAFAINSTDPAATSCWLCLSGGPPY  
EGMAKGGKFNVTKEHRNQCTWGSRNKLTLTEVSGKTCIGKAPPSHOHLCYSTVVEQ  
ASENLYLPCYNWMACTGLTPCVSTSVFNOSKDFCVWQIVPRVYHPEEVLDEY  
DYRNRPKREPVSLLTAVMLGLTAVGCTGAALITGPQOLEKGLGELHAAMTEDLR  
ALEESVSNLEESLTSLEVVLQNRGLDLLFLREGGLCAALKEECFFYVDHSGAIRDS  
MSKLRELRERRRREADQWFGWFRNRPWMTLLSALTGPLVLVLLLLTVGPCLIN  
RFVAFVRERSAVQIMVLQQOYGLLSQGETDI."

BASE COUNT 927 a 854 c 867 g 834 t

ORIGIN

alignment\_scores:

Quality: 3433.00 Length: 658

Ratio: 5.257 Gaps: 2

Percent Similarity: 99.240 Percent Identity: 98.176

alignment\_block:

US-09-171-553B-6 x PERENV2 ..

Align seg 1/1 to: PERENV2 from: 1 to: 3482

1 MetHisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17

|||||  
911 ATGCATCCACGTTTAAGCTGGCGCCACCTCCCGACTCGGGGTGGAGACC 960

17 OlysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34

|||||  
961 GAAAAGACTGAGAAATCCCTTAAGCTTCGCCCTCCATCCCTGGTTCCCTTA 1010

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50

|||||  
1011 CTCTAACAAATACTCCCGAGCCAGCTAGCTAAACGCCCTTATAGACAGCTCG 1060

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67

|||||  
1061 AACCCTCCATAGACCTTTATCCCTTACCTGGCTGATATTGACCTGATAC 1110

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84

|||||  
1111 GGGTGTCACTGTAATACACTCGAGGTGTGCTCCTAGAGGCACCTGGT 1160

84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

|||||  
1161 GGCTTGAATGCAATTCGCTCCGATTGATTAAACCCCGCTGTTAAAGC 1210

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117

|||||  
1211 ACACCTCCCAACTAGTCGTAGTTATGGGTTCTATTGCTGCCACGGCAC 1260

117 rGluLysGluLysTyrCysGlySerGlyGluSerPheCysArgArgT 134

|||||  
1261 ACAGAAAGAGAATACTGTGGGGTCTGGGAATCCTTCTGTAGGAGAT 1310

134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150

|||||  
1311 GGAGCTGGCGTCACTCCACAGATGAGACTGGAATGCCGATCTCTCTC 1360

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167

|||||  
1361 CAGGACCGGGTAAAAATTCCTTTTCTCAATTCCGCGCCGGGCAAGTACAA 1410

167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
|||||  
1411 AGTGATGAAACTATATAAGATTAAGAGCTGCTCCCCATCAGACTTAGATT 1460  
184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200  
|||||  
1461 ATCTAAAGATAAGTTTCTCACTGAAAGAGAAACAGGAAAAATATTCAAAAG 1510  
200 alaAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGlyG 216  
|||||  
1511 TGGATAAATGGTATGAGCTGGGAATAGTTTTTATAAATATGCGGGGG 1560  
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233  
|||||  
1561 AGCAGGCTCCACTTTAAACCACTTCGCTTAGATAGAGACGGGACAGAAC 1610  
233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249  
|||||  
1611 CCCCCTGGCAGTGGGACCGGATAAAGTACTGGCTGAACAGGGGCCCCCG 1660  
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266  
|||||  
1661 GCCCTGGAGCCCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGCG 1710  
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283  
|||||  
1711 GCCTGACATAACACAGCGCGCTAGCAACGGTACCCTGGATTGATTCTTA 1760  
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299  
|||||  
1761 CCAACAGCGCTAGAAACTCCCGAGGTTCCTGTTAAGACAGGACAGAGA 1810  
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316  
|||||  
1811 CTCCTTCACTCATCCAGGAGCTTTCCAAAGCCATCAACTCCACCGACCC 1860  
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333  
|||||  
1861 TGATGCCACTTCTTCTTGTGGCTTCTCTATCTCAGGGCCCTCCTTATT 1910  
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349  
|||||  
1911 ATGAGGGGATGGCTAAAGAGAAATTCAAATGTGACCAAGAGCATAGA 1960  
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366  
|||||  
1961 AATCAATGTACATGGGGTCCCGAAATTAAGCTTACCCTCACTGAAGTTTC 2010  
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 383  
|||||  
2011 CGGGAAGGGACATGCATAGSAAAGCTCCCCCATCCCAACACACTTT 2060  
383 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399  
|||||  
2061 GCTATAGTACTGTGTTTATGAGCAGCGCTCAGAAAAATCAGTATTAGTA 2110  
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416  
|||||  
2111 CCTGGTTATACAGGTGGTGGGATGCAATACTGGGTAAACCCCTGTGT 2160  
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433  
|||||  
2161 TTCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCTATGGTCAAA 2210  
433 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449  
|||||  
2211 TCGTCCCGCGAGTGTACTACCATCTCAGGAAAGTGGTCTTGTGATGATAT 2260  
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466  
|||||  
2261 GACTATCGGTATTAACCGACCAAAAGAGAACCCGCTATCCCTTACCCCTAGC 2310

```

466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 483
2311 TGTAAATGCTCGGATAGGACGGCGTGGCGTAGGAACAGGACAGCTG 2360

483 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499
2361 CCTGATCACAGGACACAGACGCTAGAGAAAGGACTTGGTAGGTACAT 2410

500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
2411 CGGGCCATCACAGAAGATCTCCGAGCCTTAGAGGAGTCTGTAGCAACCT 2460

516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgG 533
2461 AGAAGAGTCCCTGACTCTTTCTGTAAGTGGTCTACAGAACCGGAGG 2510

533 lyLeuAspLeuLeuPheLeuArgGlyGlyLeuGlyLeuGlyLeuLys 549
2511 GATTAGATCTGCTGTTCTAAGAGAGGTGGTTATGTGACGCTTAAAA 2560

550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
2561 GAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCAT 2610

566 tAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaA 583
2611 GAGCAAGCTTAGAGAAGGTTAGAGAGCGCTCGAAGGAAAGAGAGGCTG 2660

583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599
2661 ACCAGGGTGGTTTCAAGGATGTTCAACAGCTCTCCTTGGATGACCACC 2710

600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 616
2711 CTGCTTCTGCTCTCAGCGGACCCCTAGTAGTCTGCTCTGCTTACTTAC 2760

616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 633
2761 AGTGGGCTTCTGCTTAATAATAGTTTGTGGCTTTTGTAGAGAACGAG 2810

633 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
2811 TCAGTGCAGTCAGATCATGTTAGTACTTAGCAACAGTACCAAGGCTTCTG 2860

650 SerGlnGlyGluThrAspLeu 656
2861 AGCCAAGGAGAACTGACCTC 2881

```

seq\_name: gb\_vi: PEN293657

```

seq_documentation_block:
LOCUS      PEN293657      8763 bp      DNA      30-MAY-2001
DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene
            and env gene, class B, clone 213.
ACCESSION  AJ293657
VERSION    AJ293657.1 GI:14275841
KEYWORDS   env gene; envelope; gag gene; group specific antigen; pol gene;
            polymerase; polyprotein.
SOURCE     porcine endogenous type C retrovirus.
            porcine endogenous type C retrovirus.
ORGANISM   Viruses: Retroid viruses; Retroviridae; Gammaretrovirus.
REFERENCE  1 (bases 1 to 8763)
AUTHORS    Krach, U., Fischer, N., Czauderna, F. and Tonjes, R.R.
TITLE      Comparison of replication-competent molecular clones of porcine
            endogenous retrovirus class a and class b derived from pig and
            human cells
JOURNAL    J. Virol. 75 (12), 5465-5472 (2001)
MEDLINE    21256017
REFERENCE  2 (bases 1 to 8763)
AUTHORS    Tonjes, R.R.
TITLE      Direct Submission
JOURNAL    Submitted (31-JUL-2000) Tonjes R.R., Medical Biotechnology,
            Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,

```

## GERMANY

```

FEATURES             Location/Qualifiers
     source            1..8763
                     /organism="porcine endogenous type C retrovirus"
                     /proviral
                     /specific_host="Sus scrofa"
                     /db_xref="taxon:70540"
                     /clone="213"
                     /cell_line="PK15"
                     /note="class B"
     LTR               1..630
     misc_feature      467
     primer_bind       633..650
                     /note="Cap site"
     gene              1077..2651
                     /gene="gag"
                     1077..2651
                     /gene="gag"
                     /codon_start=1
                     /product="group specific antigen"
                     /protein_id="CAC39625.1"
                     /db_xref="GI:14275842"
                     /translation="MGQTVTTPLSLTLDHMTVEVRSRAHNLISVQVKKGPWOTFCASEWP
                     TFDGVPSEGTNSEIILAVKAIIFOTGPGSHPDQEPYILTWODLAEPPEPPVKKPWLNL
                     KPRKGPRIILALGKKNKHSAEKVEPSGIYPIEIPETWPEPQVPPPPYPAQGVRC
                     PSAPPKVPVEGPAAGTSRRGATPERIDEAILPRTYGPMPGQQLQPPYFFSS
                     ADLYNNKTHNHPFSDPQRLTGLVESLMFSHOPTWDDCQQLQTLFTFEERILLEA
                     RKNVPGADRGPTQLQNEIDMGFPLTRPGWDYNTAEGRESLKIYQALVAGLGSARRP
                     TNLAKREVMPQNPPEPSVFLERLMEAFRRFPDPTSEAKASVALAFIGQSALDIR
                     KKLORLEQLQAEULDLVREAKVYYRRETEEKQKKEKEREERERDRRQKNT
                     KILAAVEGKSRRERDRFKIRSPQSGNLNRTPLDKDQCAKCKGHWARNCPK
                     KNGKPKVLALEDKD"
     gene              2652..6239
                     /gene="pol"
                     <2652..6239
                     /gene="pol"
                     /note="complete pol ORF, gag-pol is expressed as a
                     polyprotein"
                     /citation=[1]
                     /codon_start=1
                     /product="polymerase"
                     /protein_id="CAC39626.1"
                     /db_xref="GI:14275843"
                     /translation="GRRGSDLPPEPRVTLKVEGQPVBEFVDTGAHSHVLQPLGKLKE
                     KKSWMGATGQRYQPTWTRTRVLDLAVGRVTHSFVLIPCEPVLPLGRDLTLTKGAQISF
                     EQGREVSNKPKITVLTLQDDERYLYSPQVKPDQDIQSWLEQFPQAAETAGMGLA
                     KQVPQVILKADSATPVSVRQPLSREAREGIMPHVQRLIQOGLVYPVQSPWNTPLLP
                     VRKGTNDYRPVQDRLREVNKRVDIHTPTVPNPYLLSALPPEDEALHDLANFRHQPC
                     RLHPTSQPLAFEMWRDPCGTGOLTWTRLPQGFKNSTPFEDEALHDLANFRHQPC
                     VTLQYVDDLLLAGATKQDCLEGTALLLESLDGLYRASAKKAQICRREVTYLYSLR
                     GGQRLTEARKTVVQIPAPITAKQVREFTAGFCRWIPGFATLAAPLPLTKKRG
                     EFSWAPHQKAPDAIKKALLSAPALAPDVTKPTLYYDERKGVARGVLTQTLGPWRR
                     PVAYLSKLDPVAGWPVCLKAIAVALIVKADKLTIGQNTIVIAHALNIVRQPP
                     DRWNTNARWTHYQSLLTERTVFAPPAALNPATLLPEETDEPTDHCHQLIIEETGYR
                     KDLDITPLTGEVLTWFTDGSYVVEGKMGAAVVDGTRTWASSLPEGTSAQKAEIM
                     ALTOALRLAEGKSINITYDSRYAFATAHVHGAIVYKQGLLTSAKELKNEIISLE
                     ALHDPKRLAIIHCPGHAKADLISGNOMADRVAKQAQAVALNPILLTIETPKAPPRQ
                     YTLDEWQEIKKIDQFSETPGCTSYSGKEITLPHKGELEYVQQLHRTLTKHLQQL
                     VRTSPYHVLRLPGVADSVYKHCVCQLVNANPSRIPFGKRLRSGKGAHWEDFTEYK
                     PAKYGNKYLIVFDVTFSGWVEAYPTKSTVAKKILEEIFFRFGIPKVIQSGNGPA
                     FVAQVSGGLAKILGIDMKLHCAYPQSSGOVERMNTIKETLTKLTETGINDWIAL
                     PFVLPFVENTPGQGLTPYELLYGGPPLVEIASVHADVLISQPLSRKALEWVRQ
                     RAKWLREAYSGEGDLQVPHRFQVGDSSVYRRHHAGNLETRWKGPLYLVLLTPTTAVKV
                     EGISTWIAHSHVKKLAPPDPSGWRAEKTEINPLKLRHLRVLVPSYNNNSPGQ"
     gene              6112..8085
                     /gene="env"
                     6112..8085
                     /gene="env"
                     /codon_start=1
                     /product="envelope"
                     /protein_id="CAC39627.1"
                     /db_xref="GI:14275844"
                     /translation="MHPTLSWRHLPTRGGEPKRLRIPLSFASIAWFLTLTITPQASSK

```

RLIDSSNPHRPLSLTLIIDPDGTGVNSTRGVAPRGTWPELHFCRLINPAVKSTP  
 PNLVRSYGFYCCPTEKEKYGSGSEPCREWSQVTSNDGDKWPISLQDRVKFSFN  
 SGPGRYKVMKLIKDKSCPSDLIDYKISFTEKGQENIKQWINGMSWGIYVKYGGN  
 GSTLIRLRIETGTEPPVAVGDKVLAEGPPALEPHNLFPQILSLURDITQPPSX  
 STTGILPTNTPRNSGVPVKTGORFLSLIQAGFAINSTDPTSSWCLSSGPPYY  
 EGMAREGFNVTKEHNRNCTWGSRNKLTLEVSFGKTCIGKAPPSHOLCNSVYVEQ  
 ASENOLVPGYNRMWACNTGLTPCVSTSVFNSKDFCVMOIIVPVVYHPEEVLDEY  
 DYRYNRKREPVSLLTAVMLGLTAVGVTGTAALITGPQOLEKGLGELHAAMTEDLR  
 ALLESVSNLESLSLSEVILQNRGLDLLPLREGGGLCAALKKECCFYVDHSGAIRS  
 MSKLRLERLRREARDOQWFGFNFSNPMWTLILLSALTGLPVVLLLLLVGVCLIN  
 RFVAFVRERVSVAQVIMVLRQYQGLLSQGETDL\*

17r  
 polyA\_signal  
 BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others  
 ORIGIN

alignment\_scores:  
 Quality: 3421.00 Length: 658  
 Ratio: 5.255 Gaps: 2  
 Percent Similarity: 98.936 Percent Identity: 98.024

## alignment\_block:

US-09-171-553B-6 x PEN293657 ..

Align seg 1/1 to: PEN293657 from: 1 to: 8763

1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGluPr 17  
 6112 ATGCATCCACGTTAAGTCGGCCACCTCCCGACTCGGGGTGGAGACC 6161  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 6162 GAAAGACTGAGATCCCTTAAGCTTCGCCCTCCATCCCTGGTTCCTTA 6211  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 6212 CTCTAACAAATAACTCCCGAGCCAGTAGTAAACGCCCTATAGACAGCTCG 6261  
 51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67  
 6262 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTACCCCTGATAC 6311  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 6312 GGGTGTCTACTGTAATAGCACTCAGGTGTGCTCTAGAGGCACCTGGT 6361  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 6362 GGCCTGAACCTGATTTCTGCTCCGATTGATTAAACCCCGCTGTTAAAGC 6411  
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 6412 ACACCTCCCAACCTAGTCCGTAGTATATGGGTCTATTGCTGCCAGGCAC 6461  
 117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT 134  
 6462 AGAGAAAGAGAAATACTGTGGGGTTCCTGGGGAATCTCTCTAGGAGAT 6511  
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 6512 GGAGCTGCGTCCACCTCCCAACGATGGAGACTGGAAATGGCCGATCTCTCTC 6561  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
 6562 CAGGACCCGGGTAATAATCTCTTTGCAATTCGGCCCGGCAAGTACAA 6611  
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
 6612 AGTGATGAACATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 6661  
 184 yrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200  
 6662 ATCTAAAGATAAGTTTCTACTGAAAAAGAAAAACAGAAATATTCAAAG 6711  
 200 alAspLysTrpTrpGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216  
 6712 TGGATAAATGGTATAGCTGGGGAATAGTTTTTTTAAATATGCGGGGG 6761  
 216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233  
 6762 AGCAGSGTCCACTTTAAACCATTCGCTTAGGATAGACGCGGACAGAAC 6811  
 233 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249  
 6812 CCCTGTGCGAGTGGAGCCGATAAAGTACTGGCTGAACAGGGGCCCCGG 6861  
 250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266  
 6862 GCCTGGAGCCACCGCATAACTTGGCGGTGCCCAATTAACCTCGCTGG 6911  
 266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283  
 6912 GCCTGACATAACACAGCGCTTAGCANCAAGTACCACCTGGATTGATTCCTA 6961  
 283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299  
 6962 CCAACAGCGCTAGAACTCCCGAGGTGTTCTCTCTCAGGGCTCTCTATT 7011  
 300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316  
 7012 CTCTTCAGTCTCATCCAGGAGCTTCCAGGCCATCACTCCACCGACC 7061  
 316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333  
 7062 TGATGCCACTTCTTCTTGTGGCTTGTCTCTCCTCAGGGCTCTCTATT 7111  
 333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349  
 7112 ATCAGGGGATGGCTAAGAAAGAAATCAATGTGACCAAGAGCATAGA 7161  
 350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSe 366  
 7162 AATCAATGTACATGGGGTCCCGAAATAAGCTTACCTCCTCAGTGAAGTTTC 7211  
 366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383  
 7212 CGGGAAGGGGACATGCATAGGAAAGCTCCCGCATCCACCAACACTTT 7261  
 383 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399  
 7262 GCAATAGTACTGTGGTTTATGACGAGCTCAGAAATCAGTATTAGTA 7311  
 400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416  
 7312 CCTGGTTAATACAGGTGGTGGCATCAATCTAGTGGTTAAACCCCTGTGT 7361  
 416 lSerThrSerValPheAsnGlnSerLysaspPheCysValMetValGlnI 433  
 7362 TTCACCTCAGTCTTCAACCAATCCAAAGATTCTCTGTGTCATGGTCCAA 7411  
 433 leValProArgValTyrThrHisProGluGluValValLeuAspGluTyr 449  
 7412 TCGTCCCCGAGTGTACTACCATCCCGAGGAGTGGTCTTGTATGATAT 7461  
 450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466  
 7462 GACTATCGGTATAACCGACCAAAAGAGAACCCGCTATCCCTTACCCCTAG 7511  
 466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483  
 7512 TGTAAATGCTCGGATTAGGGACGGCGGTGGCTAGGAACAGGACAGCTG 7561  
 483 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499  
 7562 CCCTGATCACAGGACACACAGCTAGAGAAAGGACTTGGTGAGCTACAT 7611

500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516  
|||||  
7612 GCGCCATGACAGAGATCTCCGAGCTTAGAGGAGTCTGTAGCAACCT 7661  
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533  
|||||  
7662 AGAAGAGTCCCTGACCTCTTCTGCTGAGTGGTCTACAGACCGGAGG 7711  
533 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLys 549  
|||||  
7712 GATTAGATCTGCTGTTCTTAAGAGAAGTGGGTTATGTGACGCTTAAAA 7761  
550 GluGluCysCysPheTrValaspHisSerGlyAlaIleArgAspSerMe 566  
|||||  
7762 GAGATGTTGCTCTATGTAGATCACTCAGGAGCCATCAGAGCTCCAT 7811  
566 tAsnLysLeuArgLysLeuGluArgArgArgGluArgGluGluAla 583  
|||||  
7812 GAGCAAGCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGGAAGAGAGCTG 7861  
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThr 599  
|||||  
7862 ACCAGGGGTGGTGTGAAGGATGTTCAACAGGTCTCTGGATGACCACC 7911  
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616  
|||||  
7912 CTGCTTCTGCTGACGGGGCCCCCTAGTAGTCTCTGCTTACTTAC 7961  
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633  
|||||  
7962 AGTTGGGCTTCTGTTAATTAATAGTCTTCTGCTTCTGTTAGAGACGAG 8011  
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeu 649  
|||||  
8012 TGAGTGACGTGACATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8061  
650 SerGlnGlyLeuThrAspLeu 656  
|||||  
8062 AGCCAGGAGAACTGACCTC 8082

seq\_name: gb\_v1:PERY17013

seq\_documentation\_block:  
LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000  
DEFINITION Porcine endogenous retrovirus gag, pol, env genes.  
ACCESSION Y17013  
VERSION Y17013.1 GI:3005985  
KEYWORDS env gene; gag gene; pol gene.  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.  
1 (bases 1 to 7808)  
Czauderna,F., Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.  
Establishment and characterization of molecular clones of porcine  
endogenous retroviruses replicating on human cells  
J. Virol. 74 (9), 4028-4038 (2000)  
20219394  
2 (bases 1 to 7808)  
Toenjes,R.R.  
Direct Submission  
Submitted (30-MAR-1998) R.R. Toenjes, Paul-Ehrlich Institut,  
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG

FEATURES

source  
1. .7808  
Location/Qualifiers  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"  
/tissue="kidney"  
/cell\_line="PK15"  
260. .1834  
/gene="gag"  
260. .1834

gene

CDS

/gene="gag"  
/codon\_start=1  
/protein\_id="CAA76581.1"  
/db\_xref="GI:3005986"  
/db\_xref="SPTREMBL:O73504"  
/translation="MGQVITPLSLTLDHWTVEVRSRAHNLVSQVKGKPMQPTFCASEWP  
TFDVGMPSEITNSEILAVAKIIFQTSPSHDQEPYILTWDIAEDPPVVKFWLNL  
KPKGPRILALGKHNKHAKEVSPRIYEPTEPPWPEQVPPPPYPAQAVRG  
PSAPGAVGEGPAGTTRSRGATPERTDEIAILPLRTYPPMPGQLOPLQWPFSS  
ADLYNMKTNHPFSEDPQRLTGLVESLMEHQPTWDDCOOLLOTLEAEREILLEA  
RKNPCAGDRPTOLQNEIDMGFPLRLWDYNTAGRESLKIYRQALVAGLRGASRR  
TNLAKREVQGPNEPSPFLERLMEAFRRFTFPDPTSEAKASVALAPIGOSALDIR  
KKLQRLGLEAEELRDLYREAEKAYIGRETEEEKEREEREEREERDRRQENLT  
KILAAVVEGKSSRRERDFRKIRSGPROSGNLNRTPLDKQDCAYCKEKGHWARNCPK  
KGNKGPVILALEEDKD"  
2120. .4726  
/gene="pol"  
2120. .4726  
/gene="pol"  
/codon\_start=1  
/protein\_id="CAA76582.1"  
/db\_xref="GI:3005987"  
/db\_xref="SPTREMBL:O73505"  
/translation="MGAQISFEQGREPVSNNKPTITVLTLQLDDERYLSYQVKPDQD  
IOSWLEFPQAWAETAGMLAKQVPPQVILKASATPSVSRQYPLSRAREGINPHVQ  
RLIQGILVQSPWNTPLLPVRKEGTINDYRVDLRVNRKRVQDIHPTVNPYLLS  
ALPPERNWYTLVDLDAFFCLRLHPTSQPLFAFWRDPGTGRTGTLTLPQGFKN  
PTIFDEALHDLANFRIOHPQVTLQYVDDLLLAGATKQDCEGTKALLLESLDLYR  
ASAKAQCRRREVTVGLSLRGGRHWTARKTVVQIPAPTAQKRVFELGTAGFCR  
LWIPGFATLAAPLYPLTKKEGFEFSAPEHQKAFDAIKALLSAPALAPDVTTKPTLY  
VDERGTVAGVLTQTLGPWRPVAIYLSKRLDPVAGSWPVCILKATAAVALIKVDADKL  
ETDEPETHDQHLIEETGVKRDLDIPLTGEVLTWFTDGSYVVEGKRMAGAAVVDG  
THIWASSLPCTSAQALMALTOALAEKLSINIYTDSTYAFATAHVGAIVKOR  
GLTSAGREIKNERILSLLEALHLPKRLIICPGHOKAKDLISRGNGMADRVAQOA  
AGAVNLPLIETPRAPEPRROYTLEDQWEIKKIDOFSETPGTCYTSYKELLPHKEG  
LEYVQIHLRLHLGTHLQQLVYRTPYHVRDGPVADSVVAKHVCQLVNANPSKIPP  
GKRLGSHPGAHWEVDFTEVPAKYGNKYLILVVDFTSGWVEAYPTKKTSTVYAKKY  
WRKFFRDELYLR"  
5294. .7267  
/gene="env"  
5294. .7267  
/gene="env"  
/codon\_start=1  
/protein\_id="CAA76583.1"  
/db\_xref="GI:3005988"  
/db\_xref="SPTREMBL:O73506"  
/translation="MHPITLSWRHLPTRGGEPRKLRIPLSFASIANFLTLTTPQASSK  
RLIDSSNPHRPLSLTLWLIIDPDTGVTYNSTRGVAPRGTWPFELHCLINPAVKSTP  
PNLRSYGYCCPGTEKRYCGGSEESFCRRWCVTNDGDKWPIQLQDRKFSFVN  
SGPKYKVMKLYKDKSPSDDLYLKISFTEKQENIKQWINGMSWGVYKYGGGA  
GSTLTIRLRIETGTEPPVAVGPKVLAQGGPALEPPHNPVLPOLTLRPLDITQPSN  
GTTGLIPTNTPNSGVPVKTGORLFSLIQGAFOAINSTDPDTSQCLWCLSGGPYV  
EGMAKEGFNVTKEHKNCTWGSRNKLTLEVSCKTCICGAPSHQHLCYSTVYVEQ  
ASENQYLPVGNRWACNGLTPCVSSSVFNQSCDFCMVQIVPRVYIHPHEVILDEY  
DYRNRPAREPVSILTAVMLGIGTAGVGTGAALITGPOOLEKGLGELHAAMTEDLR  
ALEPSVNLSESLTSLSEVLQNRGLDLEFLREGGLCAALKECCPYVDHSGAIRD  
MSKLRLERRRRERDOQWFGNFRSPMNTLLSALTGLPLVLLLLLTGVGCLIN  
RFVAFVRERVGAQVQIMVLRQYQGLLSQGETDL"  
BASE COUNT 2077 a 1971 c 2047 g 1713 t  
ORIGIN

alignment\_scores:

Quality: 3417.00 Length: 658  
Ratio: 5.241 Gaps: 2  
Percent Similarity: 99.088 Percent Identity: 97.720

alignment\_block:

US-09-171-553B-6 x PERY17013

Align seg 1/1 to: PERY17013 from: 1 to: 7808

```
1 MethisProThLeuSerArgArgHisLeuProThrArgGlyGlyGlnPr 17
|||||
5294 ATGCATCCACGTTAAGCTGGGCGACCTCCCGACCTCGGGGTGGAGAGCC 5343
17 GlysArgLeuArgIleProLeuSerPheAlaSerIleAlaIrrPheLeu 34
|||||
5344 GAAAGAGCTGAAGATCCCTTAAGCTTCGCCCTCCATCGCGTGGTCTCTTA 5393
34 hrLeuThrIleThrProGlnAlaSerSerIysArgLeuIleAspSer 50
|||||
5394 CTCTAACATATACTCCCGAGCGCACTAGTAACCGCTTATAGACAGCTCG 5443
51 AsnProHisArgProLeuSerLeuThrIrrPheLeuIleAspProAspTh 67
|||||
5444 AACCCCATAGACCTTTATCCCTTACCTGCGCTGATATTGACCCCTGATAC 5493
67 rGlyAlaThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
5494 GGGTGTCACTGAATATAGCACTCGAGGTTGCTCCAGAGCACTGCGT 5543
84 rPProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5544 GGCCGAGCACTGATTTCTGCTCCGATGATTAAACCGCTGTTAAAAAGC 5593
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
5594 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAGGAC 5643
117 rGlyGlyGlyLysTyrCysGlyGlySerGlyGlySerPheCysArgArgT 134
|||||
5644 AGAGAAAGAAATACCTGTGGGGTCTGAGGAATCCCTGTGTGAGAGAT 5693
134 rPSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
5694 GGAGCTGCGTCACTCCCAAGCATGGAAGCTGGAATGGCCATCTCTCTC 5743
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
|||||
5744 CAGGACCGGTAAATTCCTTGTCAATCCGCGCCGAGCAAGTCAAA 5793
167 SmetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuSpt 184
|||||
5794 AGTGTGAACCTATATTAAGATTAAGAGCTGCTCCCAACAGACTTACATT 5843
184 TyrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
|||||
5844 ATCTAAAGATTAAGTTTCTACCTGAAGAAAGAAACAGGAAATATTCAAAG 5893
200 AlaAspLysTrpTyrGlyLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
5894 TGGATTAATGGTATGAGCTGGGGAATAGTTTATTAATATATGCGGGG 5943
216 YAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlnPr 233
|||||
5944 AGCAGGCTCCCTTTAACCATTCGCCCTTAGATAGAGACGGGAGCAAGAC 5993
233 rOProValAlaMetGlyProAspLysValLeuAlaGlnGlnGlyProPro 249
|||||
5994 CCCCTGTGGAGTGGGAGACCGCATTAAGTACGCTGAACAGGGGCCCGCC 6043
250 AlaLeuGlnPrProProHisAsnLeuProValIProGlnLeuThrSerLeu 266
|||||
6044 GCCCTGGAGCCGCGCATTAACCTTCCGCTGCCCAATTAACCTCGCTGCG 6093
266 gProAspIleThrGlnPrProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
6094 GCCGACATTAACACAGCCGCTTACCAACGCTGACCTGAGATTGATTCTTA 6143
283 hrAsnThrProArgAsnSerProGlyValIProValLysThrGlyGlnArg 299
|||||
6144 CCAACAGCGCTTGAAGAACTCCCGAGGTGTCCTGTTAAGACAGACAGACAG 6193
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
|||||
|||||
6194 CTCTTCAGTGTCAATCCAGGAGGCTTTCAGACCATCAACTCCACCGACCC 6243
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333
|||||
6244 TGATGCCACTTCTTCTTGTGGCTTGTCTATCTCTCAAGGGCTTCCTTATT 6293
333 YrGlnGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArg 349
|||||
6294 ATGAGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGAGCATAGA 6343
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
|||||
6344 AATCAATGTACATGGGGGTCCCGAATTAAGCTTACCTCATCGTAAGTTTC 6393
366 rGlyLysGlyThrCysIleGlyLysAlaPrProProSerHisGlnHisLeu 383
|||||
6394 CGGGAAGGGGACATGCATAGGAAAGCTCCCATCCACCAACACCTTT 6443
383 YstYrSerThrValValIrrGlyGlnAlaSerGluAsnGlnTyrLeuVal 399
|||||
6444 GCTTATACTGTGTGTTTATGACGAGGCTCAGAAAATCAGTATTAGTA 6493
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416
|||||
6494 CTGTGTTATACAGGTGGTGGGATGCAATGCTGTTAAACCCCTGTGT 6543
416 LserThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
|||||
6544 TTCCCTCCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGTGTCCAA 6593
433 LValIProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
|||||
6594 TCGTCCCGCAGGTACTACTACATCTGAGAGAAATGCTCTTGTATGATAT 6643
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
6644 GACTATCGGTATTAACGACCAAAAGAAAGAACCCGTATCCCTTACCTTAC 6693
466 aValMetLeuGlyLeuGlyThrAlaValGlyValIrrGlyThrAlaAla 483
|||||
6694 TGTAAATGCTGGATTAGGAGCGGCTGTGGCTGAGGAAGAGGACGCTG 6743
483 LLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlnHis 499
|||||
6744 CCTGTATCAAGAGCACACAGACCTAGAGAAAGACTTGTGATGACTACAT 6793
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
|||||
6794 GCGGCATGACAGAGATCTCCGAGCTTAAAGAGAGTGTGTTAGCAACT 6843
516 uGlnGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
6844 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTTACAGAACCGGAGGG 6893
533 LysLeuAspLeuLeuPheLeuArgGluGlyLysCysAlaAlaLeuLys 549
|||||
6894 GATTATATCTGCTGTCTTAAGAGAGAGGTGATTAGTGCACCTTAAAA 6943
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
|||||
6944 GAAGAAATGTTGCTTATATAGATCACTCAGAGGACCATCAGAGACTTCAT 6993
566 LAsnLysLeuArgLysLeuGlnArgArgArgArgGluArgGluAlaAla 583
|||||
6994 GAGCAAGCTTAGAGAAAGCTTAGAGAGGCTGGAAGGAGAAAGAGGCTG 7043
583 spGlnGlyTrpPheGlnGlyTrpPheAsnArgSerProTrpMetTrpThr 599
|||||
7044 ACCAGGGGGTGTGGAAGATGTTCAACAGGTCTCTTGGATTGACACAC 7093
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
|||||
```





```

6249 AACCCCATGAGACCTTATCCCTTACTGCTGATTAATGACCCCTGATAC 6298
67 rglValAlaThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
6299 GGGGTGCTACTGTAATAGCACTGAGGTGTGCTCTAGAGCAGCCTGGT 6348
84 rpproGluLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
6349 GGCTGTGACAGCATTTCTGCTCCGATTGATTACCCGCTGTTAAAGC 6398
101 ThrProAsnLeuValAlaArgSerTrpGlyPheTrpCysCysProGly 117
6399 ACACCTCCCAACCTAGCTGCTAGTATGGTCTTATTTGCGCCAGCAGC 6448
117 rglLysGluLysTrpCysGlyLysSerGlyLysPheCysArgArgTrp 134
6449 AGGAAGAAGAAATACTGTGGGGTTCGAGCAATCCTCTGTAGAGAT 6498
134 rpsrCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
6499 GGAGCTGCGTCACCTCCCAACGATGAGACTGGAAATGGCCGATCTCTTC 6548
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
6549 CAGGACCGGGTAAATTCCTTGTCAATCCGGCCGCGCAGATACAA 6598
167 smeMetLysLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184
6599 AGTGATGAACCTATATAAGATAGAGCTGCTCCCATCAGACTTAGATT 6648
184 yrlLeuLysIleSerPheThrGluArgLysThrGlyLysTrpSerLys 200
6649 ACTTAAGATTAAGTTCTACTGAAAGGAAACAGAGAAATATCAAAAG 6698
200 alaAspLysTrpTrpGluLeuGlyAsnSerPheLeuLeuTrpGlyGly 216
6699 TGGATTAATGATGAGCTGGGGAATGTTTATTAATATGCGGGGCG 6748
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyTrpGlu 233
6749 AGAGGGTCCACTTAACCATTCGCTTAGATAGAGAGGAGGACAGAAC 6798
233 rpproValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 249
6799 CCCCTGTGGCAGTGGACCCGATTAAGTACTGTGAACAGGCGCCCGC 6848
250 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 266
6849 GCCCTGGAGCCACCGCATTAAGCTGGCGGCCCAATTACCTCGCTCG 6898
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuTrp 283
6899 GCGTACATTAACACAGCCGCTAGCAAGGATACCCCTGATGATTGCTA 6948
283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGln 299
6949 CCAACACGCTTAGAAATCCCGAGTGTCTCTTAAAGACAGACAGAGA 6998
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 316
6999 CTTTCATGTCATCCAGGAGCTTTCCAAAGCATCAACTCCAGCAGACC 7048
316 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 333
7049 TGATCCCACTTCTCTGTGGCTTGTCTATCCCTCAGGGCTCCTTATT 7098
333 yrlGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 349
7099 ATGAGGGATGCTTAAGAGAAATTCATGTGACCAAGACCATAGA 7148
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 366
|||||
7149 AATCATGTACTGCGGGTCCCGAAATAGCTTACCTCACTGAAAGTTTC 7198
366 rglLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
7199 CCGAAGAGGGACATGCATAGAAAGCTCCCATCCACCAACACCTTT 7248
383 ystYrSerThrValValTrpGluGlnAlaSerGluAsnGlnTrpLeu 399
7249 GCTATAGTACTGCTTATGAGCAGGCTCGAAGAAATCAGATATTACTA 7298
400 ProGlyTrpAsnArgTrpAlaCysAsnThrGlyLeuThrProCysVal 416
7299 CCTGTATTACAGGTGCTGGGATGCATATAGGTTTACCCCTGTGT 7348
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
7349 TTCTCTCTAGCTTCAACCAATCCAAAGATTTGTGTGATGGTCCAA 7398
433 lValProArgValTrpTrpHisProGluGluValValLeuAspGluTrp 449
7399 TCGTCCCGGAGGTACTTACCATCCTGAGGAAGTGTCTTGATGAATAT 7448
450 AspTrpArgTrpAsnArgProLysArgLysProValSerLeuThrLeu 466
7449 GACTATCGTATTAACCGCAAAAGAGAACACCGCATCTTACCTAGC 7498
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
7499 TGTATGCTCGATTAGGAGCGCGCTGGCGTAGGAACAGGACAGCTGT 7548
483 lAlaLeuThrGlyProGlnGlnLeuGlyLysGlyLeuGlnHis 499
7549 CCGTATACAGAGACACAGCAGCTAGGAAGAGCTGGGACTTCAT 7598
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsn 516
7599 GCGGCATGACAGAAAGATCTCCGAGCTTAGAGAGTCTGTACCAACT 7648
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 533
7649 AGAAGAGTCCCTGACTTCTTGTGTAAGTGTCTACAAACCGAGG 7698
533 lLeuAspLeuLeuPheLeuArgGluGlyLysCysAlaAlaLeuLys 549
7699 GATTAGATCTGCTCTTCTTAAGAGAGTGGGTATGTCCAGCTTAA 7748
550 GluLysCysCysPheTrpValAspHisSerGlyAlaIleArgAspSer 566
7749 GAAGAAATGTTGCTTCTATGATCACTCAGAGACCATCAGAGCTCAT 7798
566 tAsnLysLeuAlaArgLysLysLeuArgArgArgGluArgGluAla 583
7799 GAGCAAGCTTAGAAGAAAGTTAGAGAGCGTCGAAGGAAAGAGGCTG 7848
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThr 599
7849 ACCAGGGGTGTTAAAGATGTTCAACAGGCTCTCTTGATGATGACC 7898
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 616
7899 CTGCTTCTCTCTGAGCGGAGCCCTAGTGTGCTCTCTGTACTTAC 7948
616 rValGlyProCysLeuLeuAsnArgPheValAlaPheValArgGluArg 633
7949 AGTTGGGCTCTCTTAATTAATAGTTTGTGCTTGTAGAGAAAG 7998
633 alSerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeu 649
7999 TGAAGTGAAGTCAAGATCATGTTAGGCAACAGTACCAAGGCTTGT 8048
650 SerGlnGlyLysLeuAspLeu 656
|||||
8049 AGCCAAGAGAAACTGACCTC 8069

```



```

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
6489 ACACCTCCCAACCTAGTCGCTAGTATGGGTCTATGCTGCTGCCACGGCAC 6538
|||||
117 rGluYsgGlyTyrCysGlyGlySerGlyGlySerPheCysArgArgT 134
|||||
6539 AGAGAAAGAGAAATACTGTGGGATCTGGGGAATCTCTGTAGAGAT 6588
|||||
134 rPseCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
6589 GGAAGCTGCTCACCCTCCACCATGAGACTGGAAATGGCCGATCTCTC 6638
|||||
151 GluAspArgValLysPheSerPheValAlaSerGlyProGlyLysTyr 167
|||||
6639 CAGGACCGGGTAAATTCCTCTTGTCAATTCGCCGCCGGAAGTAA 6688
|||||
167 sMetLeuLysLeuTyrLysAspLysSerCysSerProSerAspLeuApt 184
|:::|
6689 AGTGATGAAGATATATAAGATMAAGCTGCTCCCATCAGACTAGATT 6738
|||||
184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
|||||
6739 ATCTAAGATAGTTTCACTCAAGAAAAGAAAACAGAAAATTTTCAAG 6788
|||||
200 aAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGly 216
|||||
6789 TGGATTAATGGTATGAGCTGGGGAATAGTTTATTAATATGGCGGG 6838
|||||
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
|||||
6839 ACCAGGCTCCACTTAAACCATTCGCTTAAAGTAGACAGCGGAGAC 6888
|||||
233 rOProValAlaMetGlyProAspLysValIleAlaGluGlnGlyProPro 249
|||||
6889 CCCCTGTGGCAGTGGGACCCGATTAAGTACTGCTGAACAGGGGCCCC 6938
|||||
250 aAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 266
|||||
6939 GCCCTGAGACCCACCCATTAACCTGGCGGATTAACCTGCTGCTG 6988
|||||
266 gProkAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283
|||||
6989 GCCTACATAACACAGTGCCTTACGACGATACCACTGATTTGATCTTA 7038
|||||
283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
|||||
7039 CCAACAGCCTAGAAACTCCCAAGTGTCTCTTAAAGACAGACAGACGA 7088
|||||
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316
|||||
7089 CTCTTCAGTCTCATCCAGGAGCTTTCCAGGCATCAACCTCCACGACCC 7138
|||||
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333
|||||
7139 TGAATGCCACTCTCTTGTGCTTGTCTATCCACAGGCTCTCTTATT 7188
|||||
333 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArg 349
|||||
7189 ATGAGGGATGGCTAAAGAGAGAAATTCATGTGACCAAGAGCATGA 7238
|||||
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366
|||||
7239 AATCAATGATACATGGGGGTCCGAAATTAAGCTTACCTCAGTGAAGTTTC 7288
|||||
366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
7289 CGGGAAGGGGACATCATGAGAAAGCTCCCAATCCACCAACACCTTT 7338
|||||
383 ySTyrSerThrValValTyrGlnGlnAlaSerGluAsnGlnTyrIleVal 399
|||||
7339 GCTATAGTACTGTGTATATGAGCGGCTCGAATAATCATATTATTA 7388
|||||
400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 416

```

---

```

|||||
7389 CTTGTTATTAACAGGTGGTGGCATGCATATACGGGTAAACCCCTGTGT 7438
|||||
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
|||||
7439 TTCACCTCAGTCTTCAACCAATTCACAAAGATTTCTGTGTCATGGTCCAA 7488
|||||
433 lValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
|||||
7489 TCGTCCCCGAGTGTACTACCATCTTGAGAAAGTGGCTTGTGATGAATAT 7538
|||||
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
|||||
7539 GACTATGGTATAACGACCAAAAGAGAACCCGTATCCCTTACCCCTAGC 7588
|||||
466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
7589 TGTAATGCTCGATTAGGAGCGCGCTTGGCTAGAGAACAGGAGACAGCTG 7638
|||||
483 lAlaIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 499
|||||
7639 CCTGATCAGACAGCACACAGCAGCTAGAGAAAGACTTGGTGAAGTTCAT 7688
|||||
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 516
|||||
7689 GCGGCATGACAGAAAGATCTCCGAGCCTTAGAGAGAGCTGTAGCAACT 7738
|||||
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
7739 AGAAGATCTCCCTGACTTCTTGTGTGAAGTGTCTTACACACCGGAGGG 7788
|||||
533 lLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 549
|||||
7789 GATTAGATCGTGTCTTCTTAAGAAAGTGGTATATGACGCTTAA 7838
|||||
550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 566
|||||
7839 GAAGAAATGTTGCTCTTATGATGACTCAGAGACCATCAGAGACTTCAT 7888
|||||
566 tAsnIleLeuAlaGlyLysLeuGluArgArgArgArgGluArgGlnAla 583
|:::|
7889 GAGCAACCTTAGAGAAAGTTAGAGAGCGCTGAGAGGAAAGAGGCTG 7938
|||||
583 spGlnGlyTrpPheGluGlyTyrPheAsnArgSerProTrpMetThrThr 599
|||||
7939 ACCAGGGGTGTTGAAGGATGTTCAACAGTCTCTTGGATGACACAC 7988
|||||
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
|||||
7989 CTGCTTCTGCTGACGAGACCCCTAGTACTCTGCTCTGTTACTTAC 8038
|||||
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
|||||
8039 AGTTGGGCTCTTGAATTAATAGTTTGTGCTTGTGTTAGAGAAAGAG 8088
|||||
633 aSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 649
|||||
8089 TGAATGACAGTCCAGATCATGTGTAAGTGTAGCAACAGTACCAAGGCTTCTG 8138
|||||
650 SerGlnGlyGluThrAspLeu 656
|||||
8139 AGCCAAAGAGAAACTGACCTC 8159

```

seq\_name: gb\_v1: PEN288588

seq\_documentation\_block: 2000 bp RNA VRL 29-AUG-2000

LOCUS PEN288588 2000 bp RNA VRL 29-AUG-2000

DEFINITION Porcine endogenous retrovirus PERV-AB7 env gene for envelope protein, genomic RNA.

ACCESSION AJ288588

VERSION AJ288588.1 GI:7414445

KEYWORDS env gene; envelope protein; porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus  
 Viruses; Retrod viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 2000)  
 AUTHORS Bosch, S., Arnould, C. and Jestin, A.  
 TITLE Study of full-length porcine endogenous retrovirus genomes with  
 envelope gene polymorphism in a specific-pathogen-free large white  
 swine herd

JOURNAL J. Virol. 74 (18), 8575-8581 (2000)  
 MEDLINE 20411432  
 REFERENCE 2 (bases 1 to 2000)  
 AUTHORS Bosch, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,  
 AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE

FEATURES  
 source  
 1..2000  
 /organism="porcine endogenous retrovirus"  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:61673"  
 /clone="PERV-AB7"  
 /note="French Large White SPF-pig"  
 1..606  
 /gene="env"  
 1..606  
 /gene="env"  
 /note="truncated, not functional"  
 /codon\_start=1  
 /product="envelope protein"  
 /protein\_id="CAB86207.1"  
 /db\_xref="GI:7414446"  
 /translation="MHPTLSMRHLPRTGEGPKRLRIPSPASIAWFLTLITPQASSK  
 RLIDSNPHRLSLWLIIDPTGVNSTRGVAAGTWWPELHPLRLINPAVKSTP  
 PNIVRSYGFCCPCTEKEXKCGSGESFCRMSVCVSNQDMKWPISLDQVXFSTVN  
 SGFKYKVMKLYDKSGSPDLIDYIKISTFEKKNRIIFNSG"

BASE COUNT 532 a 496 c 499 g 473 t  
 ORIGIN

alignment\_scores:  
 Quality: 3317.00 Length: 647  
 Ratio: 5.199 Gaps: 3  
 Percent Similarity: 98.609 Percent Identity: 96.291

alignment\_block:  
 US-09-171-553B-6 x PEN288588 ..

Align seg 1/1 to: PEN288588 from: 1 to: 2000

```

1 MethisProthLeuSerArgrHsiLeuProthArgrGlyGluPr 17
|||||
1 ATGCATCCACGTTAAAGCTGCGCCACCTCCGAGCTCGGGTGAGAGCC 50
17 olysArgrLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
51 GAAAGGCGTGAAGATCCCTTAAGCTTCGCCATCGCTGGTCCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerIysArgrLeuIleAspSer 50
|||||
101 CTCTAACAATAACTCCCGAGCGCACTAGTAAAGCCTTATAGACAGCTCG 150
51 AsnProHsiArgrProLeuSerLeuThrTrpLeuIleIleAspProAsp 67
|||||
151 AACCCCATAGACCTTATCCCTACCTGCGCATGATTATGACCCGTGATAC 200
67 rGlyAlaThrValAsnSerThrArgrGlyValAlaProArgrGlyThrTrp 84
|||||
201 GGGTGTCACTGTAATAGCACTCGAGGTGTCTCTTAGAGGCACTGGT 250
84 rpProGlnLeuHsiPheCysLeuArgrLeuIleAsnProAlaValIysSer 100
|||||
251 GGCCTGAACTGCATTTCTGCTCCGATGATTGAACCCGCTGTTAAAGC 300
```

```

101 ThrProAsnLeuValArgrSerTrpGlyPheTrpCysCysProGlyTh 117
|||||
301 ACACCTCCCAACCTAGTCGCGAGTATGAGGTCTATTGCTGCCAGCAGC 350
117 rGlyGlyGluIysTrpCysGlyGlySerGlyGluSerPheCysArgrTr 134
|||||
351 AGAAGAAAGAGAAATAGTGGGGGTCTGAGAGATCTCTCTGAGAGAT 400
134 rPserCysValThrSerAsnAspGlyAspTrpIysTrpProIleSerLeu 150
|||||
401 GGAGCTGCTGTCCTCCACAGATGAGAGCTGGAATATGCGATCTCTTC 450
151 GlnAspArgrValIysPheSerPheValAsnSerGlyProGlyTrpY 167
451 CAGAGCCGGGTAAATATCTCTTGTCTAATTCGCGCCGGGCAAGTACA 500
167 smetLeuIysLeuTrpIysAspIysSerCysSerProSerAspLeuAsp 184
|||||
501 AGTATGAACCTATATTAAGATAAGACTGCTCCCATCAGACTTAGATT 550
184 yrLeuIysIleSerPheThrGluArgIysThrGlyIysTrpSerIysVa 200
|||||
551 ATCTAAAGATTAAGTTTCACTGAAGAAAGCAAGAAATATTCAAAGT 600
200 LAspIysTrpTrpGlyLeuGlnIysAsnSerPheLeuLeuTrpGlyGly 216
|||||
601 GGATTAATGGATGAGCTGGGGAATAGTTTATTAATATGCGGGGGA 650
217 AlaGlySerThrLeuThrIleArgrLeuArgIleGluThrGlyThrGlu 233
|||||
651 ACAAGGTCACCTTAACCATTCGCTTAGATGAGACAGGAGAGAGACC 700
233 oProValAlaMetGlyProAspIysValIleAlaGlnGlnGlyProPro 250
|||||
701 CCTGTGCACTGGAGACCCCATTAAGTACTGCTGAACAGAGGCCCCGG 750
250 IaleuGluProProHsiAsnLeuProValProGlnLeuThrSerLeuArg 266
|||||
751 CCTGGAACCAACCGCATTAACCTGCGGGTGCCCAATTAACCTGCTGG 800
267 ProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuPro 283
|||||
801 CCTACATACACACGCGCTAGCAAGGATGACACGATGATGATTCCTAC 850
283 rAsnThrProArgrAsnSerProGlyValProValIysThrGlyIleArg 300
|||||
851 CAACAGCCCTAGAAACTCCCGAGGTGCTCTGTTAAGACAGAGACAGAA 900
300 eupHeserLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 316
|||||
901 TCTTCACTTCATCCACAGGAGGCTTCCAGACCACTCCACAGCAGCC 950
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 333
|||||
951 TGATGCCACTTCCTTGTGGCTTGTCTATCTCTCAGGGCCCTCTATT 1000
333 yrGlnGlyMetAlaIysGlnArgIysPheAsnValThrIysGlnHsiAr 349
|||||
1001 ATGAGGGGAGTGGCTAAAGAAAGAAATTCAAATGACCAAGCAAGCTA 1050
1001 ATGAGGGGAGTGGCTAAAGAAAGAAATTCAAATGACCAAGCAAGCTA 1050
350 AsnGlnCysThrTrpGlySerArgrAsnIysLeuThrLeuThrGlnVal 366
|||||
1051 AATCAATGATCATAGGGGCTCCGAAATATTAACCTCACTCAATATTC 1100
366 rGlyIysGlyThrCysIleGlyIysAlaProProSerHsiGlnHsiLeu 383
|||||
1101 CGGGAAGGAGCATGCAATAGAAAGCACTCCCATCCACCAACACTTAT 1150
383 ysrTrpSerThrValIleValTrpGlnAlaSerGlnAsnGlnTrpLeuVal 399
|||||
1151 GCTATATCTACTGTGTTATGACAGGCTCCAGAAAATCAGTATTTAGTA 1200
400 ProGlyTrpAsnArgrTrpAlaCysAsnThrGlyLeuThrProCysVa 416
```





Viruses: Retrovirdae: Retroviridae: Mammalian type C  
retroviruses: 1-Mammalian type C virus group.

REFERENCE  
1 (bases 1 to 2000)  
Bosch, S., Arnould, C. and Jestin, A.  
Study of full-length porcine endogenous retrovirus genomes with  
envelope gene polymorphism in a specific-pathogen-free large white  
swine herd  
J. Virol. 74 (18), 8575-8581 (2000)

JOURNAL  
MEDLINE  
20411432  
2 (bases 1 to 2000)  
Bosch, S., Arnould, C. and Jestin, A.  
Study of full-length porcine endogenous retroviral (PERV) genomes  
with envelope gene polymorphism in a large white  
specific-pathogen-free swine herd  
Unpublished  
3 (bases 1 to 2000)

REFERENCE  
Bosch, S.  
Direct Submision  
Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,  
AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE

FEATURES  
source  
1..2000  
/organism="porcine endogenous retrovirus"  
/specific\_host="Sus scrofa"  
/db\_xref="taxon:61673"  
/clone="PERV-B7"  
/note="French Large White SPF-pig"  
1..1899  
/gene="env"  
1..1899  
/gene="env"  
/codon\_start=1  
/product="envelope protein"  
/protein\_id="CAB86211.1"  
/db\_xref="GI:7414454"  
/translation="MHPITLSWRMLPTRGGERKRLRIPLSEFASFWMLTTLTIPOASSK  
RLIDSSNPHRLSLTWLIDPDTGVNASTGVAPRGTWPELHFCRLINPAKSTP  
PNLVRVYCCPGTEKEKCYGGSGSEFCRMSCVTENDGMMKPILODRKFSFN  
SGPGRYVMKLYKDKSCSPDLVLTISTEKGKQENIOMKINSMGIYEVKGGGA  
GSTVTRIRITGTEPPYAVGPKVLAAGCPALPEPHNLVPLDTSLRDIDQPPN  
GTTGLIPTNMRNSPVPVKTGKRLFSLIQAFRAISTDEANSSCWLCISGPPY  
EGMAKEGFVNTKEHRNCTWGSRNKLTLEVSQKGTICAPSHOHPCTSTVYEO  
ASENOYLPVGNWACNTGLPCVSTSVFNQSDFCVMOIVPVYVYHDEEVLDY  
DYRNRPKGPVSLTAVMLGLTAVGTTALITGPOOLEKGLGLHNAAMTEDR  
ALEESVNLSESLTSGVNLONRGDLPLTRGGICAALEKCCFYVDSGAIKRS  
MSKLRERLERRRRREADQWFEQEMFNRPMTLLSALGPLVVLILLTLVGCILIN  
RFVAFVRE"

BASE COUNT 532 a 485 c 505 g 477 t 1 others  
ORIGIN

alignment\_scores:  
Quality: 3278.00 Length: 646  
Ratio: 5.179 Gaps: 2  
Percent Similarity: 97.988 Percent Identity: 95.511

alignment\_block:  
US-09-171-553B-6 x PEN288592 ..

Align seg 1/1 to: PEN288592 from: 1 to: 2000

```

1 MethisProthleuSerArGArHisleuProthArglyglyIupr 17
|||||
1 ATGATGCCAAGTTAAGCTGCGCACCTTCGACCTGGGGTGAGAGCG 50
|||||
17 oLysArgleuArglleProleuSerPheAlaSerlleAlaTrPheleut 34
|||||
51 GAAAAGACTGAGAAATCCCTTAACCTTCGCTGCTGCTGCTTA 100
|||||
34 hrleuThrleThProGlnAlaSerSerlySargleuileasPserSer 50
|||||
101 CTCTAACAAATATTCCTCCAGGCCAGTAGTAACGCCCTTATAGACAGCTCG 150

```

```

51 AsnProHisArgProleuSerleuThrTrpIleuileleasProaspTh 67
|||||
151 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTTATGACCCCTGATAC 200
|||||
67 rGlyValThrValAsnSerThrArglyValAlaProArglyThrTrp 84
|||||
201 GGGTGTCTACTGTAATAGACACTGAGGTGTGCTCTAGAGGACCTGGT 250
|||||
84 rPProGluLeuHisPheCysleuArgleuileasProAlaValLysSer 100
|||||
251 GGCCTGAACCTCATTTTCCTCCGATTTGATTAAACCCCGCTGTAAGAC 300
|||||
101 ThrProAsnLeuValArgSerlyrGlyPheThrCysCysProGlyTh 117
|||||
301 ACACCTCCCAACCTAGTCCGATTAAGGCTTATGCTGACCCAGGACAC 350
|||||
117 rGluLysGluLysThrCysGlyGlySerGlyLysPheCysArgTrg 134
|||||
351 AGAGAAAGAGAAATAGTGGGGGTCTGGGGGAATCCTGTGTAGAGAT 400
|||||
134 rPserCysValThrSerAsnAspLysTrpLysTrpProileSerLeu 150
|||||
401 GGAGCTGCTGCTACCTCCACAGCATGAGACCTGGAATGGCCGATCTTCTC 450
|||||
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
|||||
451 CAGGACCGGTAAATTCCTCTTGTCAATTCGCGCCCGGCAAGTACAG 500
|||||
167 sMetMetLysLeuThrLysAspLysSerCysSerProSerAspLeu 184
|||||
501 AGTGATGAACATATATTAAGATAGAGCTGCTCCCATCAACACTTAGATT 550
|||||
184 rYLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTrpSerLys 200
|||||
551 ATCTAAGATAGATTTCACCTGAAAGAAAGAAAGAAATATTCATAAG 600
|||||
200 aLAspLysTrpTrpGlyLeuGlyLysSerPheLeu..LeuTrpGlyGly 216
|||||
601 TGGATTAATGTGATGAGCTGGGGAATGTTTTTATAAATGTGGCGGGG 650
|||||
216 yAlaGlySerThrleuThrIleArgLeuArgIleGluThrGlyThrGlu 233
|||||
651 ACAGAGGTCACCTGTAACCATTCGCTTATGATAGACAGGGGACAGAAC 700
|||||
233 rOpProValAlaMetGlyProAspLysValLeuAlaGluGluGlyPro 249
|||||
701 CCCCTGTGGCAGTGGAGCCCGCATTAAGTACGTAACAGAGGCGCCCG 750
|||||
250 AlAlaGluProProHisAsnLeuProValProGlnIleuThrSerLeu 266
|||||
751 GCCCTGGAGCCACCGCATTAATTCGACGTCGCTGTTAAGACAGAAAGAA 800
|||||
266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuilePro 283
|||||
801 GCTTGACATTAACACAGCCCTAGCAGCGTACACCTGATTTGATTCCTA 850
|||||
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 299
|||||
851 CCAACATGCTAGAACTCCCAAGGTGCTGTTAAGACAGAGAAAGAA 900
|||||
300 leuPheSerLeuIleGlyValAlaPheGlnAlaIleAsnSerThrAsp 316
|||||
901 GCTTTTCAGTCAATCCAGGAGCTTCAAGCAATCAATCAACCCAGACCC 950
|||||
316 oAspAlaThrSerSerCysTrpLeuLysLeuSerSerGlyProProTrp 333
|||||
951 GAGAGGCAATCTCTTGTGTTGCTTGTCTATCTCAGGCGCTCTTATT 1000
|||||
333 rYrGlyLysMetAlaLysGlyLysPheAsnValThrLysLysLys 349
|||||
1001 ATGACGGAGTGTGCTAAAGAAAGAAATTCATGTAACCAAGAGATAGA 1050
|||||
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrIleuThrGluVal 366

```



```

1051 AATCATATGACATGGGGGTCCGAAATTAAGCTTACCCCTACGGAATTTTC 1100
366 rgllysglythrCysAlleglyLysAlaProProserHisGlnHisLeuc 383
1101 CGGGAAGGAGACATGATAGAAAACCTCCCATCCCAACCAACACCTT 1150
383 ystfserthrValValItyrGluGlnAlaSerGluAsnGlnItyrLeuVal 399
1151 GCTATAGTACTGTGTTTATGAACAGGCTCAGAAAATCATGATTATAGTA 1200
400 ProGlyTyrAsnArtrPrpAlaCysAsnThrGlyLeuThrProCysVal 416
1201 CCTGTTATACAGCTGTGGGATGCACTACTGGGTTAACCCCTGTGT 1250
416 lserThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433
1251 TTCACCTCAGTCTTTAACCAATCCAAAGATTCTGTGTCAGTGCACAA 1300
433 lavalProArgValItyrTyrHisProGluGlnValValLeuAspGluTyr 449
1301 TCGTCCCCCGAGTACTACATCCTCGAGGAGTGTCTTGATGAATAT 1350
450 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 466
1351 GACTATCGGTATACCGACCGAAGAGAACCCGTATCCCTTACCTAGC 1400
466 aValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrGlyThrAla 483
1401 TGTATGCTCGGATTAGGAGCGCGCTTGGCGTAGAAGAGAGACAGCTG 1450
483 lalLeuIlethrGlyProGlnGlnLeuGlyGlyLeuGlyGluLeuHis 499
1451 CCTGATACACAGCACACACAGCTAGAGAAAGACTGTGAGCTACAT 1500
500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 516
1501 GCGGCATGACAGAAAGATCTCCGACCTTAGAGAGTGTGTAGCAACT 1550
516 uGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
1551 AGAAGAGTCCCTGACTCTTGTCTGGAGTGTCTACAGAACCGGAGG 1600
533 lylLeuAspLeuLeuPheLeuArgGlyGlyLeuCysAlaAlaLeuLys 549
1601 GATTAGATCTCTGTTCTTAAGAGAGGTGGTTATGTGCAGCCTTAAA 1650
550 GluGluCysCysPheTyrValAspHisSerGlyValAlaIleArgAspSer 566
1651 GAAGATGTTCTCTATGATGATCCTCAGAGCCATCAGAGACTCCAT 1700
566 tAsnLysLeuArgLysLysLeuGluArgArgArgLysArgLysAla 583
1701 GAGCAACCTTAGAGAAAGTTAGAGAGCGCTGAAAGGAAAGAGAGCTG 1750
583 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTyrMetThrThr 599
1751 ACCAGGGGTGGTTGAAGGATGTTCAACAGGTCTCCTTGATGACACACC 1800
600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 616
1801 CTGCTTCTGCTGTGACGGGACCGCTAGTAGTCTGCTCTGTTACTTAC 1850
616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
1851 AGTTGGGCTTGTCTTAATTAATAGTTTGTCCCTTGTAGAGAAATGAG 1900
633 alSerAlaValGlnIleMetValLeuArgGln 644
1901 TGAGTGCAGTCCAGATCATGTGTTAGACACAGC 1935
seq_name: gb_v1.PEN288590

```

```

seq_documentation_block:
LOCUS       PEN288590      1996 bp      RNA                VRL      29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B4 env gene for envelope
              protein, genomic RNA.
ACCESSION   AJ288590
VERSION     AJ288590.1   GI:7414449
KEYWORDS    env gene; envelope protein.
SOURCE      porcine endogenous retrovirus.
ORGANISM    porcine endogenous retrovirus.
             Viruses; Retroviral viruses; Retroviridae; Mammalian type C
REFERENCE   1 (bases 1 to 1996)
AUTHORS     Bosch, S., Arnald, C. and Juelin, A.
TITLE       Study of full-length porcine endogenous retrovirus genomes with
             envelope gene polymorphism in a specific pathogen-free large white
             swine herd
JOURNAL     J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE     20411432
REFERENCE   2 (bases 1 to 1996)
AUTHORS     Bosch, S.
TITLE       Direct Submission
JOURNAL     Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
             AFSSA-Zoopole, Zoopole des Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES
             Location/Qualifiers
             1..1996
                /organism="porcine endogenous retrovirus"
                /specific host="Sus scrofa"
                /db_xref="taxon:61673"
                /clone="PERV-B4"
                /note="French Large white SPF-pig"
             1..1215
                /gene="env"
             1..1215
                /gene="env"
             /note="truncated, not functional"
             /codon_start=1
             /product="envelope protein"
             /protein_id="CAB86209.1"
             /db_xref="GI:7414450"
             /translation="MRRPSSRRHPTKGGPKRIRIPLSFASIAWELTLVTTPQASK
             RLIISSDPHRLSLTWLIDPDVTYVNSTGVAPRGTWMELEFLRLINPAKSTP
             PNLVRSYGFCCPPTKEKYCGGSRSCGWSCVTSNDMDKWPISLDPRVRSFVN
             GSPGRKYKMWELYKDKSCSPDLDYKISFTEKRGKGNLQKWINMSGIVFYKGGGA
             GSTLIRLRIETGHTPVAVPNIVLAEDEPPALEPHNIPVPLQTLRDPDTPPN
             STGGLIPNTPRNSGVVVKGORLFLIGARQALISTDPDASSSCWLCSSPPY
             EGMAREGFWNTKRRHQCTMGSNKLTILVDSKGTICGAKAPPSHOLCNSTVIEQ
             ALEMDYLVPGTN"
BASE COUNT   521 a      499 c      503 g      472 t      1 others
ORIGIN
alignment_scores:
Quality: 3108.50      Length: 647
Ratio: 5.022          Gaps: 4
Percent Similarity: 95.672      Percent Identity: 91.345
alignment_block:
US-09-171-553b-6 x PEN288590
Align seg 1/1 to: PEN288590 from: 1 to: 1996
1 MethISProThrLeuSerArgArgHisLeuProThrArgGlyGlyLupr 17
||||| 1 ATGCGTCCACGTCACAGCGCGGCCACCTCCGACTCGGGGAGAGACC 50
17 oLysArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
||||| 51 GAAAGACAGAGAAATCCCTTAAGCTTCGCTCCATCGCTGTTCTCTTA 100
34 hLeuThrThrLeuThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
||||| 101 CTCTAACAGTGTACCTCCCAAGGCGCAGTAGTAGACGCTTATGACAGCTG 150

```

```

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAlaPth 67
151 GACCCCATAGACCTTATCCCTACGGGCTGATTATGACCCGATAC 200
67 rglValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
201 GGGTGTCACTGTTATAGCACTCGAGGTGTGCTCTAGAGGACACTGT 250
84 rpproGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
251 GCCCTGAACGATTCCTGCTCCGATGATTAACCCGCTGTAAAAAC 300
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
301 AACACCCCAACCTAGTCCGATGATGAGTTCTATGCTGCCAGGCAC 350
117 rglValSerGlyTyrCysGlyGlySerGlyLysSerPheCysArgArgT 134
351 AGAGAAAGAGAACTACTGTGGGGTTCCAGGGAATCCTGTGGGAGAT 400
134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGCTCACTCCAAAGATGGAGACTGGAATGGCCGATCTCTTC 450
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
451 CMGAGACGGGTAATAATTCTCTTGTCAATTCGGGCCCGGCGAAGTACA 500
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
501 AGTATGAGAACTATATTAAGATAGAGCTGCTCCCATCATGACTTAGAT 550
184 yrlLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200
551 ATTTAAAGATAAAGTTTCACTGAAAAAGAAACAGGGAATATATCAAAAG 600
200 alAspLysTrpTyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGlyG 216
601 TGGATTAATGATGATGAGCTGGGGAATGTTTATTAATATGAGAGGGG 650
216 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233
651 AGCAGGGTCCACTTAACATTCGCTTAGATAGAGAGCGGGCGACGGAA 700
233 roProValAlaMetGlyProAspLysValLeuIleGluGlyProPro 249
701 CCCCTGGGCAATGGGACCCACATAGTACTGGCTGAACAGAGGCCCGG 750
250 AlAlaGluProProHisAsnLeuProValProGluLeuThrSerLeuAr 266
751 GCCCTGGAGCCACACACAACTTCCGCGTCCCACTTAACCTCGCTGGG 800
266 gProAspIleThrGluProProSerAsnSerThrThrGlyLeuIleProT 283
801 GCCTGACATTAACACAGCGCGCTGACAAACATACCTGAGTTCATCTTA 850
283 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 299
851 CCAACAGGCTAGAACTCCCGGTGTTCTCTGTTAAGACAGACAGAGA 900
300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspP 316
901 CTCTTAGTCTCATCCAGGAGCTTTCAGCCATCACTCAACGACGACCC 950
316 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 333
951 CGAGCGCAAGTCTTCTGTTGGCTTGTCTATCTCAGGGCTCTCTATT 1000
333 yrlGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349
1001 ATGAGGGAATGTGCTAGAGAAAGGAATTCATATGTGACCAAGGACATAGA 1050
350 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 366

```

---

```

1051 GATCAATGTACATGGGGGCTCCGAAATAAAGCTACCTCACTAGAGTTTC 1100
366 rglValGlyGlyThrCysIleGlyLysAlaProProSerHisGlnIleu 383
1101 CGGGAAGGGGACATGATAGGAAAGCTCCCAATCCACACACACTTT 1150
383 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 399
1151 GCATAGTACATAGTATTATGAGCAGGCTTAGAAATCAGTATTAGTA 1200
400 ProGlyTyrAsnArgTyrTrpAlaCysAsnThrGlyLeuThrProCysVa 416
1201 CCTGTTATTAACGTAGTGGGCAATGCAATACCTGGTTAATCCTCTGT 1250
416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGln 433
1251 TTCACCAACAGTCTTAAACCAATCCAAAGATTCTGTGTCATAGTCCAGC 1300
433 lValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 449
1301 TCGTCCCTGGGTACACATACCATCTGAGGAAGTACCTCGATGAGTTT 1350
450 AspTyrArgTyrAsnArg..ProLysArgGluProValSerLeuThrLeu 466
1351 GACTATCGGCTACCGCTCCAAAGAGACCCCTGACCTTAACCTAG 1400
466 lValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrGlyAla 482
1401 CCGTTATACAGATAGGATAGGATGGCT..GGTGGGGGACAGGACAGCT 1447
483 AlAlaLeuIleThrGlyProGlnGluLeuGluLysGlyLeuGluLeuH 499
1448 GCCCTGATCACAGGACACACAGCTAGAAAAAGAGACTGGTGAAGTACA 1497
499 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsn 516
1498 TGCCGCCATTAACAGAAATCTCCAGGCCCTAGAGAAATCTGTAAACC 1547
516 euGluGluSerLeuThrSerLeuSerGluValValLeuGluAsnArgArg 532
1548 TAGAAGAAATCCCTGATTTCTTGTCTGTAAGGTGTCTACAACTGGAGG 1597
532 GlyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 549
1598 GGATTAGATCTGCTTCTTAATAAGAGGTGGTATGTGACGCTTAAG 1647
549 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 566
1648 ACAAGAAATGTGCTTCTATGTGATCACTCAGGAGCAATCGAGACTCCA 1697
566 eFAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 582
1698 TAGACAGCTTAGAAGAAAGTTAGAGGAGCGTGAAGGAAAGAGAGCT 1747
582 AspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrTh 599
1748 GACCAAGGGGTGTTGAAGAGATGTTCAACAGCTCTCTTGGATACACC 1797
599 rleuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuT 616
1798 CCGCTTGTGCTGTGATGGAGCTCTAGTGCTCTCTCTCTGTTGCTTA 1847
616 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 632
1848 CAGTTGGGCTTGCTTAATTATAGTGTGTTGCTTGTGTAGAGAAGA 1897
633 ValSerAlaValAlaIleMetValLeuArgGlnGln 644
1898 GTGAGTGCAGTCCAGATCATGTGACTTACACAA 1933

```

seq\_name: gb\_v1:PEN288591



```
366  ergIyLysGIYThrCysIIleGIYLeuAlaProPserHisGlnHisLeu 382
|||||
1101  CCGGGAAAGGACATGCATAGGAAAGCTCCCATCCACCAACACCTT 1150
|||||
383  CysTyrSerThrValIValTYrGIUlnAlaSerGIuAsnGlnTYrLeuVa 399
|||
1151  TGCAATAGTACTATGTTTATAGCAGGCCCTTAGAAAATCAGTATTAGT 1200
|||||
399  lProGIYTyraSnaRgtPTrPaIaCysAsnThGIYLeuThProCysV 416
|||||
1201  ACCTGGTATAACTAGATGGTGGCATGCAATACTGGGTAACTCCCTGTG 1250
|||||
416  aISerThSerValPheAsnGlnSerLysaspPheCysValMetValGln 432
|||||
1251  TTCTACACAGCTCTTACCAATCCAAAGATTCTG.GTCATGGTCCAG 1299
|||||
433  lIleValProArgValTYrTYrHisProGIUlnValValLeuAspGIUty 449
:::
1300  CTCGTCCTCCCTGGTACACTACCATCCTGAGGAGAAATCATGCATGAGTA 1349
|||||
449  rAsPTyArGTyraSnaRgtProLYsArgGIUProValSerLeuThLeuA 466
|||||
1350  GGACTATCGGCTAACCCGCTCAAAAGAGACCCGTCAGCCTAACCCCTAG 1399
|||||
466  lAValMetLeuGIYLeuGIYThrAlaValGIYThrGIYThrAla 482
|||||
1400  CCGTTATATCTAGGATTAGGATGGCT..GGTGGGGACAGGACAGCT 1446
|||||
483  AlAlaLeuIleThrGIYProGIUlnLeuGIuLysGIYLeuGIYLeuHI 499
|||||
1447  GCCCTGATCACAGACACAGCAGCTAGAAAAGACTGGTGAGCTACA 1496
|||||
499  sAlAlaMetThrGIUAspLeuArgAlaLeuLYsGIUSerValSerAsnL 516
|||||
1497  TCGCGCCATAACAGAGATCTCCGAGCCTTAGAGGAATCTGTTAACACC 1546
|||||
516  euGIUlnSerLeuThrSerLeuSerGIUValValLeuGlnAsnArgArg 532
|||||
1547  TAGAAGAAATCCCTGACTCT.TTGTCTGAGATGGTTCTACAGAACTGGAGG 1595
|||||
533  GIYLeuAspLeuLeuPheLeuArgGIUlyGIYLeuCysAlaAlaLeuLY 549
|||||
1596  GGATTAGATCTGCTGTTCTTAAAGAAAGGTGGTTATGTGCAGCCTTAAG 1645
|||||
549  sGIUlnCysCysPheTYrValAspHisSerGIYAlaIleArgAspSerM 566
:::
1646  AGAAGAAATGTCTTCTATGTGATCACTCAGAGCCATCAGAGACTCCA 1695
|||||
566  eTAsnLYsLeuArgLYsLeuGIUArgArgArgGIUArgGIUAla 582
|||||
1696  TGAGCAACCTTAGAAGAAAGTTAGAGACCTCCAAAGGGAGAGAGAGCT 1745
|||||
583  AspGlnGIYTrpPheGIUlyTYrPheAsnArgSerProTrpMetTrhTh 599
|||||
1746  GACCAGGGGTGGTTTGAAGAGATGTTCAACAGGTCTCTTGATGACCAC 1795
|||||
599  rLeuLeuSerAlaLeuThrGIYProLeuValValLeuLeuLeuLeuLeuT 616
|||||
1796  CCTGCTTCTCTCTGATGGAGACCCCTAGTGCTCTGCTCTGTTGCTTA 1845
|||||
616  hrValGIYProCysLeuIleAsnArgPheValAlaPheValArgGIUArg 632
|||||
1846  CAGTTGGGCTTGGCTTAATTAATAGGTTATTGCTTTGTTAGAGAAATGA 1895
|||||
633  ValSerAlaValAlaGlnIleMetValLeuArgGlnGln 644
|||||
1896  GTGAGTGCAGTCCAGATCATGTACTTGTAGACACGC 1931
|||||
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:53:17 ; Search time 10411.9 Seconds

(without alignments)  
4683.669 Million cell updates/sec

Title: US-09-171-553B-9

Perfect score: 2956  
Sequence: 1 tgcctttaggttagaac.....aaaaaaaaaaaaaaaaaa 2956

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenBank1:  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htg\_hum:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_inv:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2899.6	98.1	8750	14	PEN133818	AJ133818 Porcine e
2	2896.8	98.0	8196	6	A65552	A65552 Sequence 2
3	2896.4	98.0	8209	6	A65553	A65553 Sequence 3
4	2893.8	97.9	8763	14	PEN293657	AJ293657 Porcine e
5	2893.4	97.7	3482	6	AX002804	AX002804 Sequence
6	2889.4	97.7	3482	14	PEN17013	Y12239 Porcine end
7	2839.6	96.1	7808	14	PEN133816	Y17013 Porcine end
8	2811.6	95.1	8918	14	PEN133817	AJ133816 Porcine e
9	2004.4	67.8	8849	14	PEN133817	AJ133817 Porcine e
10	1970.8	66.7	1974	6	A65559	A65559 Sequence 9
11	1941.6	65.7	1971	14	AF014162	AF014162 Porcine e
12	1882.8	63.7	7333	4	AF038601	AF038601 Sus scrof
13	1882.8	63.7	7333	6	AR130474	AR130474 Sequence
14	1866.6	63.2	2000	14	PEN288592	AJ288592 Porcine e
15	1868.4	63.2	2000	14	PEN288588	AJ288588 Porcine e
16	1854	62.7	1968	14	PEN288590	AJ288589 Porcine e
17	1735.8	58.7	1996	14	PEN288590	AJ288590 Porcine e
18	1696.2	57.4	1999	14	PEN288591	AJ288587 Porcine e
19	1695.4	57.4	2002	14	PEN288591	AJ288591 Porcine e
20	1661.6	56.2	2008	14	PEN288586	AJ288586 Porcine e
21	1563.8	52.9	8918	14	PEN293656	AJ293656 Porcine e
22	1563.4	52.9	2462	6	AX002802	AX002802 Sequence
23	1552.4	52.9	2462	14	PEN288591	Y12238 Porcine end
24	1552.2	52.5	4918	6	AX052637	AX052637 Sequence
25	1490.8	50.4	8132	4	AF038600	AF038600 Sus scrof
26	1489.2	50.4	8132	6	AR130475	AR130475 Sequence
27	1483.8	50.2	8132	4	AF038599	AF038599 Sus scrof
28	1433.8	48.5	7873	6	AX052638	AX052638 Sequence
29	1231.8	41.7	8060	6	AR130473	AR130473 Sequence
30	1053.2	35.6	1996	14	PEN288584	AJ288584 Porcine e
31	1009.8	34.2	1980	6	AX052633	AX052633 Sequence
32	1009.8	34.2	1980	14	AF130444	AF130444 Porcine e
33	1007.4	34.1	2009	14	PEN288585	AJ288585 Porcine e
34	987.8	33.4	1840	4	AF296168	AF296168 Sus scrof
35	719	24.3	7362	6	AX052634	AX052634 Sequence
36	622.8	21.1	4402	6	AX052635	AX052635 Sequence
37	600.2	20.3	3320	6	A65551	A65551 Sequence 1
38	545.2	18.4	1042	14	PEN17012	Y17012 Porcine end
39	519.4	17.6	158455	2	AC079471	AC079471 Mus muscu
40	485.2	16.4	8379	14	GL060065	U60065 Gibbon leu
41	478.4	16.2	3025	4	AF147808	AF147808 Sus scrof
42	466.6	15.8	8431	4	AF151794	AF151794 Phascolar
43	460.8	15.6	590	14	PEN298074	AJ298074 Porcine e
44	455	15.4	6312	6	A60211	A60211 Sequence 7
45	455	15.4	6312	6	AR122287	AR122287 Sequence

## ALIGNMENTS

RESULT	1
LOCUS	PEN133818 8750 bp DNA
DEFINITION	Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class B, clone 43).
ACCESSION	AJ133818
VERSION	AJ133818.1 GI:6688949
KEYWORDS	env gene; env protein; gag gene; gag protein; pol gene; pol protein.
SOURCE	porcine endogenous retrovirus.
ORGANISM	porcine endogenous retrovirus
REFERENCE	Viruses; Retrovirdae; Mammalian type C
AUTHORS	retroviruses; 1 Mammalian type C virus group.
TITLE	(bases 1 to 8750)
JOURNAL	Czudeerna,F., Fischer,N., Bolter,K., Kurth,R. and Tonjes,R.R.
MEDLINE	Establishment and characterization of molecular clones of porcine
REFERENCE	endogenous retroviruses replicating on human cells
	J. Virol. 74 (9), 4028-4038 (2000)
	20219394
	2 (bases 1 to 8750)





Db	6516	AACGATGAGACTGGAAATGGCCGATCTCTCCAGAGCCGGTAAATTCCTTTGTC	6575
Qy	841	aattccggccgggcaagtagtaaaatgataaaatataaagaagctctccca	900
Db	6576	AATTCCGGCCGGGCAAGTACAAAGTGTATTAATTAAGATTAAGACTCTCCCA	6635
Qy	901	tcagactatgattatcaaaagatlaagtttcaactgaaagaagaacagaaatattca	960
Db	6636	TCAGACTTATGATTATCTAAAGATTAAGTTTCACTGAAAGGAAACGAGAAATATTCA	6695
Qy	961	aagtgtataatgtatgagctgaggaaatgttttataataatgaggaggaagga	1020
Db	6696	AAGTGTATAATGTATGAGCTGGGGAATGTTTTTATTAATATGGCCGGGAGAGAGG	6755
Qy	1021	tcacacttaacactgctctgaagatagagagcgggaacaaacccctgtggaatgga	1080
Db	6756	TCACCTTTAACATTCCCTTTAGATTAAGACAGGGGACAAACCCCTGTGGAGATGGGA	6815
Qy	1081	cccagataaagtaactgtagaacagggcccccggccctgtagccacagcaactgtgcg	1140
Db	6816	CCCGATAAAGTACTGGCTGAACAAGGGCCCGCCGCTGGAGCCACCGCATTACTTGCGC	6875
Qy	1141	gtgccecaataaactctgctgctgagctgacataacacagccgcttagcaagtaacct	1200
Db	6876	GTGCCCAATTAACCTGCTGCGGCTGACATTAACACAGCGGCTTAGCAACGGTACCCT	6935
Qy	1201	ggattgtctctcaacacagccttagaagaactcccaagtgcttctgttaagacagga	1260
Db	6936	GGATTGATTCTTACCAACAGCGCTAGAAACTCCCGAGGTTCCTGTTAAGACAGAGAG	6995
Qy	1261	agactctcagctcactccacagggagcttccaaagccactccacacagccctgtgac	1320
Db	6996	AGACTCTTCAAGTCTCATCCAGGGAGCTTCCAAAGCCATCAATCCACGACCGCTGATGCC	7055
Qy	1321	actctctctgttgcttctctctctcctcagggcctccctataatgagggatgagctaa	1380
Db	7056	ACTTCTTCTGTGGCTTGTGTCTTAATCCAGGGCTCTTATTAATGAGGGATGGCTAAA	7115
Qy	1381	gaaagaatcaatgctgacaaagaagacagaaatcaatgataatgaggggtcccaaat	1440
Db	7116	GAAAGAAATTCATGTGACCAAAAGACATTAAGAAATCAATGTACATGGGGTCCCAAT	7175
Qy	1441	aagcttaacccactgtagatgtccgggaagggagacatgacatagaaagctcccccac	1500
Db	7176	AAGCTTAACCTCACTGAAGTTTCCGGAGGGACATGATAGGAAAGCTCCCCCATCC	7235
Qy	1501	caacaaacacttctctatgactgctgtgtttagagcagggccacagaatcaatattta	1560
Db	7236	CACCAACAACCTTGTCTATGACTGTGTATGACAGGCTTCAGAAAAATCAGTATTTA	7295
Qy	1561	gtactctgttatacaggtgtgagcagatgcaatactggttlaacccctgtgtttccac	1620
Db	7296	GTACTCTGTTATACAGGTGTGTGATGCAATTAATCTGGTTAACCCCTGTGTTTCTCTC	7355
Qy	1621	tcagcttcaacaaatccaaagattgtgtgtcattggttccaaatcgtcccccagtgatc	1680
Db	7356	TCAGTCTTCAACCAATCAAAAGATTCTGTGTATGTGTCATAATTCGCCCCGAGTGTAC	7415
Qy	1681	tacacatctgaggaagtgtgtctctgtatgataatgactatcggtatataccgaaacaa	1740
Db	7416	TACCATCTCTGAGAAAGTGTCTCTGATGAATATGACTATGATGATTAACCCACCAAAAAGA	7475
Qy	1741	gaaccgtatcccttacccttagctgtatgctgagatlaaggaagggcgtgtgctgtaga	1800
Db	7476	GAACTCGTATCTTACCTTAGCTGTATGCTCGGATTAGGGAGGGCGTTGGCGTAGGA	7535
Qy	1801	acaaggacagctgcccgtatcacagacacagcagctagagaagaagactgtgtgacta	1860
Db	7536	ACAGGAGACGTGCGCTGATACAGAGCAACAGCAGCTAGAGAAAGACTTGGTGACTTA	7595
Qy	1861	catcgcgcatgacagaagatctccagacctlaaaggagctcgtttagcaactagaagag	1920
Db	7596	CATCGCGCATGACAGAAAGATCTCCGAGCCTTAGAGGAGTCTGTTAGCAACCTAGAGAG	7655
Qy	1921	tcactgactcttctgtcagaatgggttctacagaacccggaagggaatagatctgcttt	1980
Db	7656	TCCCTGACTCTTTGTCTTAAGTGTCTTACAGAAACCGAGGGATGTAGATCTGTGTTT	7715
Qy	1981	ctaagaagaagtgtgtatgtcgaacttaaaagaagaatgttctctatgataaac	2040
Db	7716	CTAAGAGAGAGTGGGTATGTGTCACCTTTAAAGAAAGATGTGCTTCTATGTAGATCAC	7775
Qy	2041	tcagaagccataagaagactccaatgaacaaagcttaaaaaaagttagaagagctcgaag	2100
Db	7776	TCAGAGCCCATAGAGACTCCATGAGCAAGCTTAAAGAAAGGTTAGAGGGGCTGAAAGG	7835
Qy	2101	gaaagaagctcagcaaggggtgtgttgaagatgttcaaaaggctcctctgtatgac	2160
Db	7836	GAAAGAGAGGCTGACAGGGGTGTGTTAAAGATGTTCAACAGGTCTCTTGATGAC	7895
Qy	2161	accctgtctctgtctcagcggggcccttagtctcgtctcgttacttaagttgg	2220
Db	7896	ACCTGTCTTCTGCTCTGACGGGACCCCTAGTAGTCTGCTCTTACTTACATTTGGG	7955
Qy	2221	ccctgttaattaatagttgtgttgccttctttagaagaaggttagtgcaatccagatc	2280
Db	7956	CCCTGTAAATTAATAGTTGTTGTGCTTTGTTAGAGAAAGTGAAGTGAAGTCCAGATC	8015
Qy	2281	atgttacttagcaacagtaaccaagcccttctgagccaaaggaaactgacttagct	2340
Db	8016	ATGTGTTCTTAGGCAACAGTACCAAGGCTTTGAGCAAGGAAAGTACCTTAGGCT	8075
Qy	2341	tcacagcttaagatagaactattacaagaagaagaatgagggaatgaagaatga	2400
Db	8076	TCCCACTTCTTAGATTAAGTAACTTAACAAAGCAAGAACTGGGGAATGAAGAGTAA	8135
Qy	2401	tccaacttaacccctcccaagaaccaggaagttaataaagaagcttcaatgcccgaat	2460
Db	8136	TGCAACTTAACCTCTCCCAAGACCCAGAGAGTTAATTAAGAGCTTAATGCCCGGAAT	8195
Qy	2461	acagaacctgtcgtgctcagtaaatgaatgaaggttcaacttctacttgttccagggc	2520
Db	8196	CCAGACCTGTGCTGCTGCCAGTAAATAGGTAGAGGTCACTCTTATGTTCCAGGGC	8255
Qy	2521	ctgtactcctgtgctcagtaagaataacaggaatgaatgttacttaactcgttgcatt	2580
Db	8256	CTGTCTTCTGTGGCTTAAGTAAGATTAACAGGAATGATGACTTAATCGCTTATCGATT	8315
Qy	2581	ctgtaaactgactggtccacatagaagaatltatcaacttgaacagcccttagacta	2640
Db	8316	CTGTAAACTGACTGACCATTAAGAAATTAATTAACATTGACAGCCCTTAGTACTTA	8375
Qy	2641	tctcaactgcaatctgtcaactctgcccagaagccacagcagatgtcgaacctcggagcta	2700
Db	8376	TCTCAACTGCATTCGTTCATCTGTGCCAGAGGCCACGAGATGGGAGCTCCGGAGCTA	8435
Qy	2701	ttttaaatgattgtgtccaggaagcgggctctcgtatattttaaatagaattgtccatg	2760
Db	8436	TTTTAAATAGATTGTGTCCAGCGGCGGCTCTCATTTTAAATATATTGTGTCACAG	8495
Qy	2761	gagcggcgctcgtcgtatattttaaatagaattgtgtgaacgaaggttggttgtaga	2820
Db	8496	GAGCGGGGCTCTCGATTATTTAAATGATGTGTTGTGACGACAGGCTTGTGTGTGA	8555
Qy	2821	cccataaagaactgtcccatctccgactcggggcccgagctcctactacccctgtggtg	2880
Db	8556	CCCATTAAGAGCTGTCCGATTCGACATTCGAGGCGCGGAGTCTTACCCCTGCGGTGG	8615
Qy	2881	tacgactgtggcccccagcgtctgtgaataaaatcctcttgcgtgtttgtaacaaa	2938
Db	8616	TACGACTGTGGGCGCCAGCGGCTTGGAATAAATCTCTTGTGTTGTGATCAAGA	8673
RESULT	2		
A66552			
LOCUS	A66552	8196 bp	DNA PAT 29-MAR-1999



Db 7040 ACAGGGACAGCTGCCATGATCACAGACCAAGACAGCTAGAGAAAGACTTGGTACCTA 7099  
 QY 1861 catggcccatgacagaagaatctccgagccttaagaagctctgttagcaactagaagag 1920  
 Db 7100 CATCGGCGCATGACAGAAAGATCTCCGAGCCCTTAAAGAGAGTCTGTATGCAACCTAAGAGAG 7159  
 QY 1921 tccctgactctctgtctgaagtggtcttcacagaaccggagggatagatgcctgttt 1980  
 Db 7160 TCCTTGACTTCTTGTCTGAAGTGTCTTACAGACCGGAGGGATTAAGTTCGTCTTT 7219  
 QY 1981 ctaagaagaagtggtgtatctgcaagccttaagaagaatgtgtctctctatgtatagac 2040  
 Db 7220 CTAAGAGAAGGTGGGTATGTGACGCTTAAAGAAAGATGTGCTTCTATGTATGATGATCAC 7279  
 QY 2041 tcaagaagcctcagaagctcatgaagaacttagaagaagaatgtatagaagagcgcgaag 2100  
 Db 7280 TCAGGACCCATCAGAGACTCCATGACAACTTAGAAGAAAGATTAGAGAGCGCTCGAAGG 7339  
 QY 2101 gaaagaagaagctgacacaggggtgttgaagagatgtgtcaacaggtcctctgtgagac 2160  
 Db 7340 GAAAGAGAAGGCTGACAGAGGGGTGTTGAAGGATGTTCAACAGGTCTCCTTGATGATGACC 7399  
 QY 2161 accctgtctctgctctgacggggcccttagtagtccctgctcctgtacttaacagtggg 2220  
 Db 7400 ACCCTGCTTCTGCTGACGGGGCCCTAGTAGTCTCTCTCTGTTACTTACAGTTGGG 7459  
 QY 2221 ccttgacttaataagtggtgtgtccttgtagagaagaatgtatgagtcagtcacagac 2280  
 Db 7460 CTTGCTTAATTAATTAAGTGTGTTGCTTGTAGAACAAAGATGAGTGCAGTCCAGATC 7519  
 QY 2281 atgtacttaggcaacagatcccaagggcctctgacccaagaagaactgaacctagacct 2340  
 Db 7520 ATGTTACTTAGGCAACAGTACCAAGGCTTCTGAGCCAGGACAAACTGACCTTGACCT 7579  
 QY 2341 tccagctctcaagattagaactatcaacaagaagaatgtgggaaatgaaagatgaaaa 2400  
 Db 7580 TCCAGATTCTAAGATTAGAACTATTAAACAGACAAAGAGTGGGAATGAAGAGATGAAAA 7639  
 QY 2401 tgaacctaaacctcccaagaaccagaagaatgataaagaagctctaagtgccccgaat 2460  
 Db 7640 TGCAACTTAACCTCCAGAACCCAGGAAGTAAATAAAAAGCTCTAATATGCCCCCAAT 7699  
 QY 2461 aacagacctgctgctccagtaagatagagaagtgcaacctctatgttcagggc 2520  
 Db 7700 MCAGACCTCTGCTGCTGCCGTAATAGTAGAAGGTACACTTCTTATTTGTTCCAGGGC 7759  
 QY 2521 ctgctatccctgacctaaagtaagaatacagaagaatgagttagtaactcgtctatctgact 2580  
 Db 7760 CTGCTATCCTGGCCTAAGTAGTAACAGGAATGAGTTGACTAATCGCTTATCTGAT 7819  
 QY 2581 ctgtlaaactgactgacacacatagaagaatgtatcaatctgacagccctagtgacta 2640  
 Db 7820 CTGTAAAACTGACTGGACCATAGAAAGAAATGTATACATATGACAGCCCTATGTGACCTA 7879  
 QY 2641 tctcaactgaaactctgaactctgcccagaagaccccgagatgagagacctccgagacta 2700  
 Db 7880 TCTCACTGCAATCTGTGCACTCTGCCAGAGAGCCCAAGATGCGGACTCCGGAGCTA 7939  
 QY 2701 ttttaaaatgattgtgtccacagagcgaggctctgacatatttaaaatgattgtgtcatg 2760  
 Db 7940 TTTTAAATATGATTGTGTCACAGAGCGGGCTCTGATATTTTAAATGATTGTGTCATG 7999  
 QY 2761 gaagcgaggctctcgaatatttaaaatgattgtgtgacagcagagcttggtgtgaa 2820  
 Db 8000 GAGCGGGGCTCTCGATATTTTAAATGATTGTGTTGTAACGACAGGCTTGTCTTGTA 8059  
 QY 2821 ccccataaagctgtcccgatccgacactcgagcgagcgagctcctaacccctgcgctggtg 2880  
 Db 8060 CCCCATAAAGCTGTCCGATTCGACACTCGGGCGGACGACTCTTACCCCTGCGTGTG 8119  
 QY 2881 taagactgtggcccccagcgagcttgaaataaactcctctgtgctgttgacatcaaaaaa 2940

Db 8120 TACGACTGTGGGCCCGACGCGCTTGATTAATAATCTCTGTGTTGATCAACAAAA 8179  
 QY 2941 aaaaaaaaaaaaaa 2956  
 Db 8180 AAAAAAAAAAAAAA 8195

RESULT 3  
 A66553  
 LOCUS A66553  
 DEFINITION Sequence 3 from Patent WO9740167.  
 ACCESSION A66553  
 VERSION A66553.1  
 KEYWORDS GI:4538106  
 SOURCE  
 ORGANISM  
 porcine endogenous retrovirus.  
 porcine endogenous retrovirus  
 Viruses; Retroid viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 (bases 1 to 8209)  
 Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.  
 Patent: WO 9740167-A 3 30-OCT-1997;  
 Q ONE BIOTECH LTD (GB)  
 Location/Qualifiers  
 1..8209

FEATURES  
 source  
 1..8209  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 588..2162  
 /codon\_start=1  
 /product="VIRAL CORE PROTEIN"  
 /protein\_id="CAB39347.1"  
 /db\_xref="GI:4538107"  
 /translation="MGQIVTTPSLTLDHTVEVSRKHNLSVQKKKPMWTFCASENP  
 TFDVGPBSEGTENSEITIIAVALIFQTGPGSHPDQEPVLTWQDLADDPWAKPWN  
 KPRKGPRIILALGKKNKISAKRKPSSSYLPDRGADLADGTPCSEPTPLSTGCCG  
 TSAPGAPVAPVPAAGTGRSRGATPERIDEALILPLTTPGPMGGQLOLOVYFSS  
 ADLNMKTNHPPFSEDPORLTGLVLSMFESHQPMWDCQOLQTLPTTBEEERILTPA  
 RKNVAGARPTOLQNETDMQFPPLTRGQMDVNTAGDPSLKIYQALVAGLRGSRP  
 TNLAKRVNMGDPNPPSPVLERLMEAFRRPTTPDPPTSEAKKASVALAFIGOSALDJR  
 KILQRLBGLQAEALRDVLEAEKVYRRETEERDQREKREERERERRORRQKRLT  
 KILAAVVGKSSRRERDFRIRSGPROSGNLGNRTPLDKQCAVCKEKGHWANCPK  
 KGNKPKVLALEEDKD"  
 <2163..5747  
 /codon\_start=1  
 /product="REVERSE TRANSCRIPTASE"  
 /protein\_id="CAB39348.1"  
 /db\_xref="GI:4538108"  
 /translation="GRGSDPLPPRYTLKVEGQVPEFLVDGAEHSVLLOPLKKE  
 KSNMGATGGRQYPMTRRTVLDVGVGRVTHSPVTPCEPVPLIGRLDLTKMAQISF  
 BOGPEYSVNNKPTTVLTLOIDDEYRLYSQVKKPDQDIOISLBOFPDAMETAGMGIA  
 KQPPVOYLKASATPVSVKQYPLSREARBSIMHVORLIQOGLVIVQSPMNPPLP  
 VRKPTNDYRPVQDLREVRKQVODIHPTVPKPNLISALPERKMYVVDLKAQFCL  
 RLHPTSQPLFAEMRDPGTGTGQTLTTRLPQGNKSPITFDEALHRDLANFRIOHQ  
 VTLLOVYDDLILAGATKODCLEGRKALLETSLDGLYRASAKAOICREYTVLGYSLR  
 GGQRLTEARKKIVVOIIPATTAQVPEPSTGAGFCFLMIPGATLAAPLYPLTKKG  
 GFSAPRHOAKAFDILKRALISAPLAPDVTKEPTLYVDERKVARVLTOTIGPWR  
 PVATLSKLDPVAGGWPVCLKALIAVALIVKADKLTIGONTITVIAHALENITVROP  
 DRWMTNRMHTYVLSLLTERTEAPRALINPATLIPETDEPVTIHDHLLIEBTGR  
 KDLDPLTGEVILVMTDSSSYVEGRKMAAAYVDGRTIMASLSLEPSTSAQAEIM  
 ALHGPALAEKSSINITYDSYAPATAGVATIKQGLATLSAGRELTKNEEILISLE  
 ALHPLRLAITHCGHOKANDLISRGNDMDRAVKAQAOAVNLPLITFEPKAPBP  
 VYLEDQMEIKKIQDFSEPTGCTGYSGKTLPLKELIYVQOIRHTLHGTGKHLQD  
 VRTSPYVILRPVADSVVHKVPCQVLVNNPSTIPGKRLRSGHPRGAEVDETEK  
 PAKYGNKYLLVFTVESGWAEVPTKRETSVAKKLTLEIPFPFGIPKVIQSDNGFA  
 FVAQVSOGLAKILGIDMKLHCAYRPOSSGQVGRNRTIKETLKLTTETRIINMVAL  
 PFLVLEFRNTPPGGILPPEYLVLGGPPLAEIIFAHSAVDVLLQPLTSRLKALEWVO  
 RAMKOLREAYSGDLOVPHREOVDSYVVRHRRGNLETWKRGVYLITPTTAVAVE  
 GIPWIIASHVHKAPPPDSQMRKENTENPLKLRHLRILVPSNNNSGQ"  
 5620..7590  
 /codon\_start=1  
 /product="VIRAL ENVELOPE PROTEIN"  
 /protein\_id="CAB39349.1"  
 /db\_xref="GI:4538109"

CDS

/translation="MNPRLSRRLPTRGEPKRLRPLSPASIAMFLTLTTPPOASK  
RLIDSNRPRLSLWLIDPDTGVNSTRGVAAPGTWMLPCLRLINAVSTP  
PLIDRSYFCCPCEKTEKYGSGSGSFRMSCTGMDKMLSLDRLPVSFVN  
SGGKXKMMKLYKDKSPSDIDYKLTSEKTEKYSVYDKWLSNLLYGGGAG  
STTTTLRIETGTEPPVAPKPVLAEGPPALPBPNNLPVQYLSLRDITOPNS  
TGLIPLNTPRNSPGVPVATGDRFSLIGAFQALINSDPDATSSCWCLSGPPEYE  
GNKERRFPVNTKEHNOCTWGSRNKLTLEVSGKICIGKAPSHOHLICSTVVEQA  
SEMOYLVPYCNRMWACNIGLPCVSTVSNQSDKVMQIVPRVYHPEEVLDEYD  
YRNRPKREPVLTLAVMLGTAAGVGTGALITGPOLKEGCEYDHAATEDLRA  
LKESVNLSESLTSEVYLNORIGLIDFLREGGLCAALKKEGCEYDHAATEDSM  
NKLKLEPERRRRRRREADQCFEGMFRSRMNTLSALTLGPLVLLLTIVPCLINR  
FAFVVERYSAYQWYLRKOOTGLSOGSETDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t  
ORIGIN

Query Match 98.0%; Score 2896.4; DB 6; Length 8209;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2939; Conservative 0; Mismatches 11; Indels 6; Gaps 3;

QY 1 tgcctttaaggttaagaaaccccttgagacgtttgagctgacccctatgaattgctc 60  
DB 5260 TGCCTTTTAAAGGTGAGAGAACCCCTGAGACGTTGGCGTGACCCCTATGATGTGCTC 5319  
QY 61 acggggagaccccccgctggtgagaaatgctctgtaacagtgctgtagtgctctt 120  
DB 5320 ACCGGGGAGACCCCGCTTGGCAGAAATTCCTTGGACATAGCTGTGATGCTCCTCTT 5379  
QY 121 cccagccttctctcctcagcctcaagcgcctgagtggtgaggaacagagctggaag 180  
DB 5380 CCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCGAGTGGTGGAGGACCGACGCTGGAAGC 5439  
QY 181 agctcggagagcctactcagagagagagacttgcaattcacaatcgcttccaaagt 240  
DB 5440 AGCTCCGGAGGCGCTACTCAGAG--GAGACTTGCAGATCCACATCCCTTCCAAGTTG 5496  
QY 241 gagatlaagctatgtagagcgcacgcgtgcaagaaaccccgagctcgtgtagaagagc 300  
DB 5497 GAGATTAAGTCTATGTTAGAGCCACCGTGCAGAGAACTCTGAGACTCGGTGGAAGGGAC 5556  
QY 301 ctatcctgctactttgacacacacacagcgtgtgaagtcgaagaaatctccacttga 360  
DB 5557 CTATCTCGTACTTTTGACCAACACAGCGCTGTGAAGTGAAGGATGCCCACTCGA 5616  
QY 361 tcaatgcatcccaagcttaagctggtgcacactcccgactcgggtggtgagagacgaaga 420  
DB 5617 TCATGTGATCCCAAGTTAACCGCGCGCACCTCCCGACTCGGGGTGAGAGCCGAAAGA 5676  
QY 421 ctgaagaatcccttaagctcgcctccatcgctgtgtcttacttaataataatctcc 480  
DB 5677 CTGAGAAATCCCTTAACTTCGCTTCATGCGCTGGTGGTCTTACTTAACAAATTAACCTCC 5736  
QY 481 caggccagtagtaaacgcttatagacagctcgaaaccccatagacacttatccctaac 540  
DB 5737 CAGGCCAGTAGTAAGCGCTTATAGACAGCTCGAACCCCATAGACTTATTCCTTAAAC 5796  
QY 541 tgcgtgatatgaacctgatacgggtgtcaatgtaataagaactcgaagtgctgctcct 600  
DB 5797 TGGCTGATTAATTAACCTGTATAGCGGTGTCACTGTAATAAGCACTGAGAGTGTGTGCTCT 5856  
QY 601 aagagcaacctggtgacctgaactgcatcttcgctccgattgatttaaccccggtttaaa 660  
DB 5857 AGAGGACCTGTGGTGGCTGAGACTGCAATTTGCTCCGATCGAATTAACCCCGCTGTTAA 5916  
QY 661 agacacactcccaactagtcgtagtlatggttctatgtctgcgcagagcaagagaa 720  
DB 5917 AGCAGACTCTCCCAACCTAGTCGTAAGTATGAGTTCATGTGCTGCCAGGCAAGAGAAA 5976  
QY 721 gagaataactggtggtggtctctggtggaatctcttgaagagatgagctggtcacctcc 780  
DB 5977 GAGAAATTAATCTGTGGGCTTCTGAGGAACTCTTGTAGAGAGATGAGAGCTGCGTCACTCC 6036  
QY 781 aacgatgagacagtaagatggtgcgatctctctcagagaccgggttaaaatctccttgctc 840

DB 6037 AACGATGAGAGATGAGAAAGCGCATCTCTCCAGGACCGGGTAAATTTCTCCTTTGTC 6096  
QY 841 aattcggcgccgggaaggtacaaaatgtagaacatataaaagataaagctgtccccc 900  
DB 6097 AATTCGGCGCCGGGCAAGTACAAATATGATGAACATATATTAAGATTAAGCTGCTCCCA 6156  
QY 901 tcaagcttgaatatacctaagaataagattccactgaaagagaaacagaaataatcaaa 960  
DB 6157 TCAGACTTATGATTAATCTAAGATTAAGTTCACCTG--AAAGGAAACAGAGAAATTTCAA 6214  
QY 961 aagtgataaatagtatagctggtggtggaatgtttttaaataatgctgggtgagcagg 1020  
DB 6215 AAGTGATTAATGATATGATGAGCTGGGGAATAGTTTTT-ATATATAGCGGGGAGCAGG 6273  
QY 1021 tccacttaacatctgccttgaagtagagagagagagacacccctgttgcaattgga 1080  
DB 6274 TCCACTTTAACCATTTCGCTTATGATTAAGAGAGGAGGACAGAACCCCTGTGCAATGGGA 6333  
QY 1081 cccgataaagtagctgctgaacagagggcccgccctgagagccacgcatlaacttgcg 1140  
DB 6334 CCCGATAAAGTACTGTGTAACAGAGGCGCCCGGCTGAGAGCCACCGCATTAATTTGCG 6393  
QY 1141 gtgccccaatlaactcgtcgtggtggtggtggtggtggtggtggtggtggtggtggt 1200  
DB 6394 GTGCCCAATTAATCTGCTGCGGCTGATTAACACAGCGCGCTTACGACAGTACCACT 6453  
QY 1201 ggaattgattctcacaacacagcctagaaactcccaagtgcttctgtaagagagagag 1260  
DB 6454 GGATGATTTCTTACCAACAGAGCGCTAGAAATCTCCAGGTTGTTCTGTTAAGACAGAGAG 6513  
QY 1261 agactctcagctcatalcagaagagcttccaagccatcaactccacgacactgagc 1320  
DB 6514 AGACTCTTCACTCATCCAGGAGCTTCCAGGCAATCACTCCAGGCTGATGCTC 6573  
QY 1321 actctctctgtgtgctgtgtctatcccaagagcctctttaaagagagagagagag 1380  
DB 6574 ACTTCTTCTGTTGGCTTTGTCTATCTCAAGGCTCTCTTATTAAGAGGAGTGGCTTAA 6633  
QY 1381 gaaagaaatlaactgtagacaaagagacataagaaatgtagagaggttccgaaat 1440  
DB 6634 GAAAGAAATTAATAGTGAACCAAGAGCATAGAAATCAATGTAGATGGGGTCCGGAAT 6693  
QY 1441 aagcttaacctactggaagttccgggaagagagagagagagagagagagagagagag 1500  
DB 6694 AAGCTTAACCTTCACTGAAAGTTTCCGGGAAGGGAATCATAGGAAGAGCTCCCAATCC 6753  
QY 1501 caacaacaccttgcctatagtagctgtgtttagagagagcctcagaagaaatcatat 1560  
DB 6754 CACCAACACCTTGTGATATGATCTGTGTTATGAGAGGCTCGAAGAAATCAGATTTA 6813  
QY 1561 gtacctggtatlaacaggttggcagatgcaataacttgggttaacccctgtgttccacc 1620  
DB 6814 GTACCTGTTATTAACAGGAGGTGGGCAATGCAATCTGGGTTAAACCCCTGTGTTCCACC 6873  
QY 1621 tcaagcttcaacaaatcaaaagattgtgtgtcataggtccaaatctgtcccgagtgac 1680  
DB 6874 TCAGTCTTCAACCAATCAAAAGATTTCTGTGCAATGATGCAATGCTCCCGAGGTGAC 6933  
QY 1681 taccatctcgaagagagtgctcttgaatgaatatgactcgtgataacccgaaagaa 1740  
DB 6934 TACCATCTCGAGGAAGTGTCTTATGATATGATCTGCGTATTAACCGACCAAGAAAGA 6993  
QY 1741 gaaccgtaaccttaacctagctgtaaigtctcgatlaagagagcgcgtgtggtgagta 1800  
DB 6994 GAACCGTATCTCTTACCTAGCTGTATGTGCGGATTAAGGAGAGGCGGCTGGGTAGGA 7053  
QY 1801 acaggaacagctgacctgatacagagacacagcagcagcagagagagagagagagagta 1860  
DB 7054 ACAGGAGACAGCTGCCCTATATCACAGACACAGCAGCTTAGAGAAAGGACTTGTGAGCTA 7113  
QY 1861 catgctgcatgacagagagatctcagagcttaagagaggtctgttaagaaactcgaagag 1920

Db 7114 CATGGGCGCATGACAGAAATCTCCGAGCCTTAAAGAGTCTGTAGCAACCTTGAAGAG 7113  
 QY 1921 tccctgactctctgtcgaagtgtctacagaaccggaaggtatgatctgttt 1980  
 Db 7114 TCCCTGACTCTTGTGTCTGAAGTGGTTCTACAGAACCGGAGGATTAAGATCTCTGT 7223  
 QY 1981 ctaagaagaagtgtgtatgtcgaagccttaaaagaatgtgtctctatgtatgacac 2040  
 Db 7234 CTAAAGAGAGTGGTGTATGTGACAGCCTTAAAGAAAGAAATGTTCTCTATGATGATCAC 7293  
 QY 2041 tcaagaagccatcagagatcatalgaagaatgtgaaaaaagttagaagagcgctcgaag 2100  
 Db 7294 TCAGAGCGCATCAGAGACTCCATACACAGCTTAGAAAAAGTTAGAGAGCGCGGAGAG 7353  
 QY 2101 gaaagagagctgacacaggggtgtgtgaagatgtgtcacaagctcctctgtatgac 2160  
 Db 7354 GAAAGAGAGCTGACACAGGGGTGTGTGAAGATGTGTCAACAGTCTCTTGGATACC 7413  
 QY 2161 accctgtctctgtcgtcgaagggccctagtagtctgtcctgtacttaacagtgg 2220  
 Db 7414 ACCCTGCTTCTGCTGACGAGGCGCCCTAGTAGTCTGCTGCTTACTTACAGTTGGG 7473  
 QY 2221 ccttgcttaataatagttgtgtgcttgttagaagaagtgagtgagtcagtcagatc 2280  
 Db 7474 CTTTGCTTAATTAATAGGTTTGTGCTTGTGTTAGAGACAGAGTGAAGTGCAGATC 7533  
 QY 2281 atgttacttaggcaacagtagtaccagagcctctgagccaagagaagaactgaccttag 2340  
 Db 7534 ATGCTACTTAGGCAACAGTACCAAGGCTTCTGAGCCCAAGAGAACTGACCTTACCT 7593  
 QY 2341 tccagttcctaagaattagaacttaacaagaagaagtgtggaatgaaagatgaanaa 2400  
 Db 7594 TCCAGTTCTTAAGATTAAGACTATTACAGACAGAGAGTGGGAATGAAGAGTGAANA 7553  
 QY 2401 tgcgaactaaccccccaagaagccagaagatlaaataaaagccttaaatgtccccgaat 2460  
 Db 7654 TGCAGACTTAACCTCCCGAGAACCCAGGAAGTTAATAAAAAAGCTCTAATATCCCGCAAT 7713  
 QY 2461 acagaaccctgtcgtcgaatagtagtagaaggtgcaacttccatgtgtccagggc 2520  
 Db 7714 ACAGACCTGCTGCTGCTGCCAGTAAATAGTGAAGTGCACACTTCTATTTGCCAGGGC 7773  
 QY 2521 ctgtatccctggcctaaatagaataacaggaatgtgtgaactaactogttacttgat 2580  
 Db 7774 CTGTATCTGCTGCTGCTTAAGTAAATACAGGAATGAGTTGACTTAATCGCTTATCTGAT 7833  
 QY 2581 ctglaaaactgactggacacatagaagaatgtgtacacatgacagccttagtgacta 2640  
 Db 7834 CTGTAAACTGACTGGCACCATAGAAAGATTGATTACACTTTGACAGCCCTTACTGACTTA 7893  
 QY 2641 tctcaactgcaactgtcactgtcccaagagcccaagcagatgacgtcccgagacta 2700  
 Db 7894 TCTCAACTGCAATCTGTCACTTGCCAGAGAGCCCAAGAGATGGCGACTCCGGAGACTA 7953  
 QY 2701 ttttaaaatgattgtgtcccaagagcggtgtcgtatcttaaaatgattgtgtccatg 2760  
 Db 7954 TTTTAAATGATTGTGTCCAGAGCGCGGCTCTGCAATTTTAAATGATTGTGTCCATG 8013  
 QY 2761 gaagcgaggtctcgaatttttaaaatgattgtgtgacagagaggtttgtgtga 2820  
 Db 8014 GAGCGCGGCTCTCGATATTTTTAAATGATTGTGTTGTGACGACAGGCTTGTGTGTGA 8073  
 QY 2821 ccccatataaagtgtcccgatctccgactcgcgcagtcctcctacccctgtcgtgt 2880  
 Db 8074 CCCCATAAAGCTGTCCGATCTCCGACTCCGAGCGCGGAGTCTTACCCCTGCGTGGTG 8133  
 QY 2881 tagactgtgtggcccaagcggtgtggaataaaatccctgtgtgtgtgtgtgtgtgtgt 2940  
 Db 8134 TAGACTGTGTGGCCCAAGCGCGCTTGAATAAAAATCTGCTGTGTGTGATCAAAAAA 8193  
 QY 2941 aaaaaaaaaaaaaa 2956  
 Db 8194 AAAAAAAAAAAAAA 8209

RESULT 4  
 LOCUS PEN293657  
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene and env gene, class B, clone 213.  
 ACCESSION AJ293657  
 VERSION AJ293657.1 GI:14275841  
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene; polymerase; polypeptide.  
 SOURCE porcine endogenous type C retrovirus.  
 ORGANISM Porcine endogenous type C retrovirus.  
 REFERENCE 1 (bases 1 to 8763)  
 Krach, U., Fischer, N., Cauderna, F. and Tonjes, R.R. Comparison of replication-competent molecular clones of porcine endogenous retrovirus class a and class b derived from pig and human cells  
 J. Virol. 75 (12), 5465-5472 (2001)  
 JOURNAL MEDLINE 21256017  
 REFERENCE 2 (bases 1 to 8763)  
 Tonjes, R.R. Direct Submission  
 TITLE Submitted (31-JUL-2000) Tonjes R.R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225, GERMANY  
 JOURNAL  
 FEATURES  
 source  
 Location/Qualifiers  
 1..8763  
 /organism="porcine endogenous type C retrovirus"  
 /proliferal  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:70540"  
 /clone="213"  
 /cell\_line="PK15"  
 /note="Class B"  
 1..630  
 467  
 /note="Cap site"  
 633..650  
 1077..2651  
 /gene="gag"  
 1077..2651  
 /gene="gag"  
 /codon\_start=1  
 /product="group specific antigen"  
 /protein\_id="CAC39625.1"  
 /db\_xref="GI:14275842"  
 /translation="MGOTVTPPLSLTLDHMTVRSRAHNLISVYKKGPMOTPCASEMP  
 TFDVGPSEGTFFNSEIILAVKALIFQPGSHPODEYILTWDLAEPPPMVPMIN  
 KPRGSPETILALGKKNHSAKVPSPGIVPEIPEPTMPEPOVPPYPACGAVG  
 PSAPGAYVYGPRAKRSRGATPERTDELAIIPLTTGPPHGGGLQPLQTPSS  
 ADLNMKTNHPSESDPQRLGLVESLMFSHPQWDDQDLQTLFTTERERILLRA  
 RKNVYADGRETOLQNEIDMGFLPLRGWDYNTAEGRSESLKTYROALVAGLRGASRP  
 TNLAKREVMOGPMPEPVSFLERLMEAFRRTPDPPTSEAKQASVALAFGOSLDIR  
 KILQRLGLOAEFLRDLYVREAEKXYRRETEEEKQEKEREERERERRORDKNT  
 KILAAVEKSSRRERDPFRIRSGPQSGNLGRNTLDMQCAVCKEKHMARNCX  
 KNGCPVLALEFKD"  
 2652..6239  
 /gene="pol"  
 /db\_xref="GI:14275843"  
 /translation="GRGSDPLPEPRVTLKVEGOPYEFLVDTGAHSLVLLQPLGRIAKE  
 KKSVMKATGQROYPTTTRTVDLAVGRVTHSLVPECPVPLGLRLTKGAQISF  
 EOGREPVSNKPLITVLTLOLDEYRIYSQVAPDDIOISWLEQFOAANAETAGMGLA  
 KQVPOVVIQKASATPVSRQYPLSRBARBGIVPHVQRLIQDGLIVPQVSPMWTPLLP

```

VRKPGINDYRPVODIREVNRKRODHPYTPNPYNLLSALPERNNVTYDLDAFECFL
RLHPISQPLAEWRDPTGRTGQLTWLRLPGFKSPTEFDEALHRLDANRILQHPQ
VTLLOQVDDLLAGATKODLBETKALLLESLDGRASAKAOICORVATYLSLR
GGGRMLTEARKIVVOIPAPTTAKOYREELGTAGFLMIPEFATLAAYLTLEKRG
ESFMADEHOKAFDAIKKALLSAPALALPVPTKLYLVERGVAGVLTLYEMBR
PVAVLSKTLDPVAGSWPCLKAIAAVIILKADKLTGONITVIAAPHLNEMVOPR
DRMTNARKTHYOSLLTERVTFAPPAALINPALTLEEDDEVYTHDCHLLEETGVR
KDLTDLPLTEVLTWFTDSSVVEBKKAGAAVVDGTITWASSLPBETSQKAEIM
ALTOALRLAEKSNITYDSRYAFATAHVHGAIVKORGLTTSAGREIKKEEILSLE
ALHLPKRLAIHICPGHOKAKDLISNGNOADRYAKOAAQAVNLLPIETPKAPEPRQ
YLEDQOEIKKIDQSETEGECYTSYKEILLPHKGLLEYOQIHLPLTGLKHLQOL
VRTSPYHVLRLPEGVADVVKHCVPOLVANDSPRIIPGKRLRISGHAWEDEPFEVK
PAVYGNKYLIVYDFRSWVEAYPPKKESTYVAKKILIEIETPRGICPVVISDNGPA
PAAQVSGQIAKILIGIDWKLCATRPDSSQOVERMNTIETILTKLTETGINDWALP
PVLFRVNTPGQFGLTPLYELLYGPPPLVEIASVASADVLISQPLESKALEWRQ
RAWKOLREAYSGEGDLOVPHRFQVDSVVRHRAHNLTEPKNGPVLVLLTPTAVKY
EGISTWISHSHVYLAAPPDPSGWRAEKTEMLKRLHRLVPSNNNSPGQ"
6112..8085
/gene="envy"
6112..8085
/codon_start=1
/product="envelope"
/protein_id="CAC39627.1"
/db_xref="GI:14275844"
/translation="MPTLSWRHLPTRGGEPRKRIPLSPASIANPLTITTPQSSK
RLTDSNPHPUSLTWLTIDPDIGTVNSTRGVAPRGTPWPELHPLRLINAVASTP
PNLVRYGECCEGTKEKRYCGSGSESFCRMSCVTNSNGDKMWPISLQDRKSEFN
SGPKKYVMKILYKDKSCSPSDLDYLIKISPEKQKDNQIKWINGSMWIVPKYGGGA
GSTLLTIRLEIETGEPVPAVGPDKVLAEGPAPLEPHMLPVOLSLTDLTTPQPSX
STGLIPTNTPRNSGPVKTGORLESLQGAOFLNSTDPAATSCWCLSSGPYX
EGMAKKEKNVTRKEHNOCTGMSRNKLTILEVSGKTCIGKAKPSIOHLCHNTVVEO
ASBNQYLVGKNNWMAKNTGLIPCVSTSVNSKDPVCWQIVPPIRYHPEEVLIDEX
DYRNRKREPVSLTAVMLGLTAVGVTGTAALITGQJLEKGIHLEHAAETEDLR
ALEESVNLSESLTSLSEVQLNRRGLDILFLREGCLALKEECFVYDHSGLAIRDS
MSKILRELRERRERADOGWEGEFMNRPMWMTTLLSALTGLPVLVLLITVGPOLIN
REVAFVERVSAQVIMVLRQYQGLISQETDL"
8133..8763
LNR
polyA_signal 8655..8660
BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
Query Match 97.9%; Score 2893.8; DB 14; Length 8763;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2910; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 1 tgctttaggttagaacaaccccttgacagatttggtctgacccctatgaattgctc 60
Db 5749 TGCCTTTTAGGGTTAGAACACCCCTTGACAGTTTGCGCGACCCCTATGATTTGCTC 5808
QY 61 aagggggaaccccccggttgtagaagcttgctctgtaacatgctgtagctgctt 120
Db 5809 ACGGGGGACCCCCCTTGATGAAATTCCTCTGACATGCTGATGCTGCTGCTT 5868
QY 121 ccaagccttgcttcaggtcaagcgctgagtggtgaggaacacgagcgltgaagc 180
Db 5869 CCCAGCCTTTGTTCTCAGGCTCAAGCGCTCGAGTGGTGAAGCAACGAGCGTTGAGGC 5928
QY 181 agctccggaagcctactcagagagagagacttgcaagttccacatcgcttccaagt 240
Db 5929 AGCTCCGGAGGCTTACTCAGAGAAAGAGACTTGCMAAGTTCCACATCGCTTCAAGTG 5988
QY 241 gagattcagctatgttagagcgccggtgcaagaaaccccgagactcggtggaagac 300
Db 5989 GAGATTAGTCTATGTTAGAGCCACCATCATGCAAGAAACCTCGAGACTCGGTGGAAGGCC 6048
QY 301 ctatctcgtactttgaccacaacacgagctgtgaagtcgaaggaatcccaacttga 360
Db 6049 CTTATCTCGTACTTTTACACACACAAACGCTGTGAAGTCTGAGAGATTTCCACCTGGA 6108
QY 361 tccatgatcccaactgaagctgagccacactccgactcggtggtggaagacgaaaaa 420

```

```

Db 6109 TCCATGCAATCCACGTTAAAGTGGCCGACCTCCCGACTCGGGGTGGAGAGCCGAAAAA 6168
QY 421 ctgagaatcccttaagcttcgctccatcgcttgcttactacttaacaataactcc 480
Db 6169 CTGAGAAATCCCTTAAGCTTCGCTCCATCGCGCTGTTCTACTCTTAACATAACTCC 6228
QY 481 cagggcagtaagtaaacgcttataagcagcttgaaacccccataagccttaccctac 540
Db 6229 CAGGCCAGTAGTAACGCGCTTATAGACAGCTGAAACCCCAAGACCTTTATCCCTTAC 6288
QY 541 tggctatattgacccctatacaggtgtcaactgtaaaataagcagcgagtgctctc 600
Db 6289 TGGCTATATTGACCCCTGATACGGGTGTGCACTGTAATAGACCTCGAGGTGTTCTCT 6348
QY 601 aagagcacttgctggcctgaactgtaattctgctccgattgataaccgcgtgtaa 660
Db 6349 ACAGGACACTGGTGGCTGAACTGCATTTCTCCATGATTAACCCGCTGTAAA 6408
QY 661 agcacacctcccaactgctgtagtaltatgggtctatgctgcccaggaacagagaa 720
Db 6409 ACACACACTCCCAACCTAGTCCGTAGTTAGGCTTCTATGCTGCCACAGGACACAGAAA 6468
QY 721 gagaataactgtggggtcttgaggaaactctctgtaagagaatgagctgcaactcc 780
Db 6469 GAGAAATACTGTGGGGTCTTGGGGAATCCTTCTTAGAGAGATGAGCTGGCTACCTCC 6528
QY 781 aacgatgagactggaatgagcagatctctccagacccgggtaaatctccttctc 840
Db 6529 AACGATGAGGAGCTGGAAGAGGCGCATCTCTCCAGGACCGGTAAATTTCTCTTTTTC 6588
QY 841 aatccggcgccggcagagacaaatgataaacatataaagataaagagctgcctcca 900
Db 6589 AATTCGGCGCCGGGCAAGTACAAAGTGAATGAACTATTAAGATTAAGATGCTGCTCCA 6648
QY 901 tcaagattagattatcctaaagaataagttcactgtaaaaaaggaacacagaaatata 960
Db 6649 TCAGACTTATGATTACTTAAGATTAAGATTAAGTTCACTGAAAAAGGAAACAGAAAAAT 6708
QY 961 aagtgataataatgatatgagctggggaatagtttataataatagcgggggagcag 1020
Db 6709 AAGTGATTAAGATTAAGTATGAGCTGGGGAATAGTTTATTAATATAGGCGGAGCAAG 6768
QY 1021 tccacttaacatcgcccttagagatagagacgggaacacccctgtgcaatggga 1080
Db 6769 TCCACTTTAACCATTGCGCTTAGATAGAGACGGGAGACAGAACCCCTGTGGAGTGGGA 6828
QY 1081 cccgataaagtaactgctgtaacaggggcccccgccctgtagccacacgataacttgc 1140
Db 6829 CCCGATTAAGTACTGAGTGAACAGGAGGCCCCCGGCTGAGGCCACCGCATTAACCTTGG 6888
QY 1141 gtgcccaatataactcgctgctgagctgtacataaacaacgctcagtaacaagtaacc 1200
Db 6889 GTGCCCAATTAACCTCGCTGGCGCTGACATTAACACAGCGGCTTAGCANAGTACCACT 6948
QY 1201 ggaattgattctaccacaacgcttagaaactcccaagtgcttcctgttaagacagaa 1260
Db 6949 GGATTGATTCCTACCAACAGCGCTTAGAAATCCCAAGTGTTCCTGTTAAGACAGACAG 7008
QY 1261 agactcttcagtcacatccagggagcttccaagcattcaactccaacgacactgagc 1320
Db 7009 AGACTCTTCAGTCTATCCAGGAGACTTTCACAAACATCACTCCACCGACCTGATGCC 7068
QY 1321 actcttctctgttgcttcttcctcccaagggcctcctttaaaggaaggaagctaa 1380
Db 7069 ACTTCTCTTGTGGCTTGTCTATCTTAAGGCGCTCTCTTATTAAGAGGAGTGGCTAAA 7128
QY 1381 gaaagaaaaatcaatgtagcacaagaagacatagaatcaatgtacatgggggtccgaa 1440
Db 7129 GAAAGAAAATTAATGTGACCAAGACATAGAAATCAATGTAGATGGGGGTCCCGAAT 7188
QY 1441 aagcttaacctactgaagtttcggggaaggagacatgcatagaagaaagctccccaat 1500
Db 7189 AAGCTTACCTCTACTGAAGTTTCCGGGAAGGAGCATGATAGGAAAGCTCCCATCC 7248

```

QY 1501 caccacaaccttgctagtagtctgtgttatagaagagcctcaagaataatcaatttta 1560  
 Db 7249 CACCAACACCTTTGGCAAGTACTGTGTTATGAGCAGGCGCTCAGAAATCACTATTTA 7308  
 QY 1561 gtacctgtttaacaggtgtgtgagcaatagctagtttaacccctgtgttccacc 1620  
 Db 7309 GTACTGTGTTATACAGAGGTGTGGCATGCAATACTGGGTTTAAACCCCTGTGTTTCCACC 7368  
 QY 1621 tcaagcttcaacccaalccaagaattgtgttcaatgtgttcaatcgttccccaggtgtt 1680  
 Db 7369 TCAGCTTCAACCAATCCAAAGATTTCGTCTCATGTGTCACCAATCGTCCCGCAGGTGAC 7428  
 QY 1681 taacatcctgaggaagtgtgtcctgtatataactatcggtataacccaagaata 1740  
 Db 7429 TACCACTCTGAGGAAGTGTCTGATGAAATGACTATACGATTAACCCGACCAAAAGA 7488  
 QY 1741 gaaccctgataccttagcttagctgttaatgtctgattaggaagagccgtgtgtgtgga 1800  
 Db 7489 GAACCCGTATCCCTTACCTAGCTGTATGCTCGGATTAGGAGACGCGCTTGCGGTAGA 7548  
 QY 1801 acaaggaagactgtccctgtatcaacaggaacagcagctagagaagaactgtgtgtgta 1860  
 Db 7549 ACAGGAGACAGTGCCTGTATCAGACAGACAGCAGCTAGAGAAAGACTTGTGAGCTA 7608  
 QY 1861 catcgagccatgacagaagaatctccgagccttaagaaggtctgttagaaccttagaag 1920  
 Db 7609 CATGCGGCGCATACAGAAAGATCTCGAGCTTAGAGAGTCTGTAGACACCTTAGAAGAG 7668  
 QY 1921 tcccgacttctgttctgaagtgtgttctacagaacccgagaggtgtgtgtgtgtgt 1980  
 Db 7669 TCCCTGACTCTTGTGTGAAATGCTTACAGAACCGAGGCGATTGATCTGTGTGT 7728  
 QY 1981 ctaagaagaagtgtgtgtatgtgtcagccttaagaagaagtgtgtgtgtgtgtgtgt 2040  
 Db 7729 CTAAGAAGAAGT 7788  
 QY 2041 tcaagaagccatgaagactccatgaagaagcttagaagaagaagtgtgtgtgtgtgt 2100  
 Db 7789 TCAAGGAGCCATCAGAACTCCATGAGCAAGCTTAGAAGAAAGTTAGAGAGCGTGAAG 7848  
 QY 2101 gaaagaagagctgacacagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2160  
 Db 7849 GAAAGAGAGGTGACACAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7908  
 QY 2161 accgtgttctgtctgt 2220  
 Db 7909 ACCGTGTCTTGT 7968  
 QY 2221 cctgttctaataataggt 2280  
 Db 7969 CCTGTCTTATTAATAGGT 8028  
 QY 2281 atgttacttaggaagaacagagccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340  
 Db 8029 ATGTACTTAGGCAACAGTACCAAGGCTTCTGAGCCAGAGAAATGACCTTACGCT 8088  
 QY 2341 tcccaattctaagattagaactataaagaagaagaagtgtgtgtgtgtgtgtgtgtgt 2400  
 Db 8089 TCCCACTTCTAAGATTAGACTATTTAACAAGACAAAGATGGGGAATGGAAGATGAANA 8148  
 QY 2401 tgaacttaaaccttcccaagaacccaagaagttaataaagaacttaaatgtccccgaatt 2460  
 Db 8149 TGAACCTTAACCTTCCAGAACCCAGAGAAAGTTAATAAAGCTTAATGATGATGATGAT 8208  
 QY 2461 acagaccgt 2520  
 Db 8209 CCAAGACCTGT 8268  
 QY 2521 ctgtcttctgt 2580  
 Db 8269 CTGCTATCTGT 8328

QY 2581 ctgtaaactgactggcacatagaagaattgtatlaacatagtagacccctagtagccta 2640  
 Db 8329 CTGTAAACACTGACTGGCACCATAGAAAGATTGATTACACTTAGACGCCCTAGTACTGA 8388  
 QY 2641 tctcaactgtcaatctgtcaatctgtccaggaagcccaagatgtggaacctccggagcta 2700  
 Db 8389 TCTGAACCTGCAATCTGTGACTGTGCGCAGTAGGCCACGAGAGAGCGAGCTCCGAGCTA 8448  
 QY 2701 ttttaaatatgtatgt 2760  
 Db 8449 TTTTAAATATGATTGT 8508  
 QY 2761 gagcgcggt 2820  
 Db 8509 GAGCGCGGCTGT 8568  
 QY 2821 ccccatlaaagaactgt 2880  
 Db 8569 CCCCATAAAGCTGT 8628  
 QY 2881 taagactgt 2940  
 Db 8629 TACGACTGT 8686

## RESULT 5

LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000  
 DEFINITION Sequence 3 from Patent WO9853104.  
 ACCESSION AX002804  
 VERSION AX002804.1 GI:9885132

## KEYWORDS

SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus.  
 Viruses: Retroviridae: Mammalian type C  
 retroviruses: 1-Mammalian type C virus group.

## REFERENCE

1 (bases 1 to 3482)  
 Stoye, J.P. and Weiss, R.A.  
 Detection of retroviral subtypes based upon envelope specific  
 sequences  
 JOURNAL MEDICAL RES COUNCIL (GB): STOYE JONATHAN PAUL (GB)  
 LOCATION/Qualifiers  
 1. 3482  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"

## FEATURES

source  
 BASE COUNT 927 a 854 c 867 g 834 t  
 ORIGIN

Query Match 97.7%; Score 2889.4; DB 6; Length 3482;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 1 tgcgttttaggttagaagaacccctgtgagaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 60  
 Db 548 TGCCTTTTAGGTTAGGAGAACCCCTGTGACAGTTTGGCTGACCCCTATGAATTTGCTCT 607  
 QY 61 acggggagacccccccgt 120  
 Db 608 ACGGGGAGACCCCCCGT 667  
 QY 121 cccagccttgttctctaggtcagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180  
 Db 668 CCCAGCCTGTGTCTGT 727  
 QY 181 agcttcggagagagctaccagaagaagagagagctgtgtgtgtgtgtgtgtgtgtgtgt 240  
 Db 728 AGCTCGGAGAGCTTCTCAGAGAGAGAGACTTGTCAAGTTTCCATTCGCTTCCAAAGTG 787  
 QY 241 gagattcagctatgttagagccacgltgagaagaacactgtgagactcgtgtgagaaggac 300  
 Db 788 GAGATTTCAGTGTATTGTAGAGCCACCGTGCAGGAACCTGAGACTGTGTGTGTGTGTGT 847



[illegible]





Db	1388	AAATTCGGCCGGGCAAGTACAAGATGAAGAACTATTAAAGTAAAGCTGCTCCCA	1444
QY	901	tcgaacttagattatcctaagaagaagttcactgaaaaagaaacagaaataatc	960
Db	1448	TCGAGCTTGAATATCTTAAAGATTAAGTTTCCCTGAGAAAAGAAACAGAAATATTC	1507
QY	961	aagtgataaataatgtagtcgtgggaataagtttttaataatagcggggagcagg	10226
Db	1508	AAATGGTAATATGTATGAGCTGGGGAAATATGTTTTTATTAATATGCGGGGACACGG	1567
QY	1021	tccaacttaaccatcgccttagaagaagaacgggaacccccgtgtgcaatggga	10806
Db	1568	TCACATTAAACATTTGGCTTAGAGATAGAGACGGGACAGAACCCCTGTGCAATGGGA	16272
QY	1081	ccccataaagtagctggtcgtgaacaggcccccgccctggagccacgcataactgtcg	11400
Db	1628	CCGATTAAGTATGAGCTTAAACAGGGGCCCGCGCCCTTGAGACCCACATTAATCTGGCG	16877
QY	1141	gtgcccacaatlaaccctcgctgsgctgtgacataaaccgcccgttagaaacagtaacca	12000
Db	1688	GTGCCCCAATTAACTCGCTGGCGCTTGACATTAACACGCGCCTTAGACAGGTAACCACT	17474
QY	1201	ggattgattcctaaccacacgccttagaaacatccccaagttgttcgtttaaagacagag	12606
Db	1748	GGATTGATTTCTTAACCAACACAGCGCTTAGAAATCTCCACGGTGTTCCTTTAATACAGAGCAG	18077
QY	1261	agaccttcaagtcataccagaggagctttccaagccatacaactccaccgacctgtatgc	13200
Db	1808	AGACTCTTCAAGTCTATCCAGGAGGACTTCCAAAGCATCAACTCACCGACCTGATGGCC	18677
QY	1321	actctctctgtgtgcttgttctatccctcaagggtcctctattagagggaatgtc	13800
Db	1868	ACTCTCTTGTGTGGCTTGTGTATCCACAGGCGCTCCTTTATTAAGAGGGATGGCTAAA	19277
QY	1381	gaagaagaaatlaatgtgacccaagaagatagaatgaatgaatgtagtggtgtccgaat	14400
Db	1928	GAAAGAAATTTCAATGTAGACCAAAAGCATAGAAATCAATGTACTGTGGGTGCCGAAAT	19877
QY	1441	aagcttaaccctcaactgtgaagttccgggaaggagacatgcataggaanaagctcccccac	15000
Db	1988	AAGCTTACCCCTCACTGAAGTATCCGGGAAGGGAGATCATAGGAAAGCTCCCCATTC	20477
QY	1501	gaccaaacaccttgctatagtaagtggtttatagacagcctcagaanaatcagatatta	15600
Db	2048	CACCAACACCTTGTGATATAGTACTGTGTTATAGCGAGGCTCAGAAAACATATTTA	21077
QY	1561	gtaccctgttataaagaagttgtgtgcaatgcaataactgtgttaacccctgtgttcacc	16200
Db	2108	GTACTGTGTTATTAACAGGTGGGCAATGCAATACTGGTTAAACCCCTGTGTTCCACC	21677
QY	1621	tcagcttccaaccaatccaagaattgtgtgtcatgtgtccaatcgttccccagatgtac	16800
Db	2168	TCAGTCTTCAACCAATCCAAAGATTTCGTCTCATGTGCCAAATCGTCCCGCAGTATAC	22277
QY	1681	tacacatccgaagaagtgctcttgatagaataatacattcgttaacccgacccaaaaaga	17400
Db	2228	TACCATCTCGAAGAGTGTGCTCTTGATGAATATACATATGCTATTAACGACAAAAAGA	22877
QY	1741	gaacccgatacccttaaccctagctgttaaatgctcagattaggaacgcgctgtgtatgga	18000
Db	2288	GAAACCGATCCCTTACCCTAGCTAGCTAATGCTCGGATTAGGAGAGCGCGTGTGGCTATGGA	23477
QY	1801	acaaggacagctgcccctgtatcaacaagaccacagcagctlagagaaaggactgtgtagcta	18600
Db	2348	ACAGGACAGCTGCCCTGATACAGAGACCAACAGACGACTAGAGAAAGACTTGGTAGCTA	24077
QY	1861	catggggccatagacaagaagatctccggagccttaagaagttcgtltagcaacttagaag	19200
Db	2408	CATGGGCGCAATAGACGAAGATCTCCGAGCCTTAAGAGAGTCTGTTACCAACTTAAGAGAG	24677
QY	1921	tcctcgaactcttgtctgaagtggttctacaagaacccgggggattagaatgtcgtgttt	19800
Db	2468	TCCCGAATCTTTTGTCTGAAGTGGTCTCTACAGAACCGGAGGGATTAAGTATGCTGTGTT	25277

QY	1981	ctaaagaagaggggttattgtgcagcccttaaaagaagaatggtgtctcatitabagtaac	2040
Db	2528	CTAAGAGAAAGGTGGGTTATGTGCGACCTTTAAAGAAAGAAATGTTGCTCTATGTAGATCAC	2587
QY	2041	tcaagagccatcaagagatccatibgaacagcttagaanaaaaglttagagagcgctcgaag	2100
Db	2588	TCAGAGGCATCAGAGACTCCATGAGCAGCTTAAAGAAAGTTAAGAGGGCTGGAAG	2644
QY	2101	gaagaagagcgctgaccaggggtgtttgaaagatbgtltcaacagltctccttgabgac	2160
Db	2648	GAAAGAGAGGCTGACCAAGGGGTGGTTGAAGGATGTTCAACAGGCTCTTGATGATGACC	2707
QY	2161	accgcgtcttcgtcctbaagggccccttagtagtgcctctcctgttacttaagttgg	2220
Db	2708	ACCCGTGCTTTCGCTCTGACGGAGCCCTTAGTAGTGTCTGCTCTGTTACTTACAATTGG	2767
QY	2221	cctgtcctaattaaataggttgtgtgccttgtgtagaagacgaagtgaigtcaatccaa	2280
Db	2768	CCCTGCTTAATTAATAGGTGTGTTGCTTGTTAGAAGACGATGTGATCCACATGC	2827
QY	2281	atggtacttaagcaacggtaccaaagcctctgaagccaaagagaaactgacctagaacct	2340
Db	2828	ATGTGACTTTAGGCAACAGATCAAGGGCTTCTGAGCAAGAGAAATGACCTCTAGACCT	2887
QY	2341	tcccaagttctaagattgaacatattacaagaagaagttgggaaatgaagaatgaa	2400
Db	2888	TCCCGAGTTCTAAGATTAGAACATTATTAACAAGACAAAGATGGGGANTGAAGATGA	2947
QY	2401	tgcacactcaacccctcccaagaccccaagaaatlaataaaagctctaaatgcccagaa	2460
Db	2948	TGCACACTTAACCCCTCCCAAGAACCCAGAAATTAATAAAGCTTAATAATGCCCGGAATT	3007
QY	2461	acagaacctgtgtgctgtccagtaaatagttagaaggtcaaccttccatttgttccaggc	2520
Db	3008	CCAGACCCCTGTGCTGTCCAGTAATAAGTAGAAGGTCAACATTCCTATTGTTCAGAGGC	3067
QY	2521	ctgtatctctgacctaaagtaagaataacagaagaaatbgaattgactaaatgcttatctgatt	2580
Db	3068	CTGCTATCTCTGGCTTAAGTAGAATTAACAGGAATGATGTTGACTTAATGCTTATCTGGATT	3127
QY	2581	ctgtanaaactgactgtgcacacataagaagaattgattacacatlgacagccctagtgacta	2640
Db	3128	CTGTAAAAACCGACACGACCATAGAGAAATTTGATTACATGTAGACAGCCCTTAGGACTTA	3187
QY	2641	tctaacctgaaatctgtacatctgtgcccaaggagcccaagatgagacctccggagcta	2700
Db	3188	TCTACACTGCAATGTGTCACTCTGCCCCAGAGGCCACAGAGATGGGGACACCTCCGGAACCTA	3247
QY	2701	ttttaaaatgatbgtlccacagagagcgcgcgccctcgatattttaaaatgattgtgtccatg	2760
Db	3248	TTTTAAATGATGTGTCCACGAGAGCGCGGCTCGATATTTTAAATGATTTGGTTCACG	3307
QY	2761	gagcgcgggctc-tcgatatttaaaatgattggttctgtgacgcagagcttgttga	2819
Db	3308	GAGGCGGGGCTTTCGATTTTAAATTAATATTGTTGTGACGCGCACAGGCTTTGTGTGA	3367
QY	2820	accocataaagatgtcccgatcccgatctggagcgagttcctcttaacccctgtcggt	2879
Db	3368	ACCCCATTAAGCTGTCCGATTTCCGACTTCGGACTCCGGGCCCCAGCTCTTAACCCCTCGTGGT	3427
QY	2880	gtacagactgtgggcccacagcgcgcttgagataaaaactcctctgtgtttgcatc	2934
Db	3428	GTAGACTGTGGGGCCCAAGCGGCTTGGAATATAAAATCTCTGTGCTGTGGTTCATC	3482
RESULT 7			
PERV17013	PERV17013	7808 bp	DNA
LOCUS	Porcine endogenous retrovirus gag, pol, env genes.	VRL	14-APR-2000
DEFINITION	Y17013		
ACCESSION	Y17013.1	GI:3005985	
KEYWORDS	env gene; gag gene; pol gene.		

SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
REFERENCE Viruses; Retroviral viruses; Mammalian type C  
AUTHORS 1 (bases 1 to 7808)  
TITLE Cauderna, F., Fischer, N., Bolter, K., Kurth, R. and Tonjes, R.R.  
JOURNAL Establishment and characterization of molecular clones of porcine  
MEDLINE endogenous retrovirus replicating on human cells  
20219394  
REFERENCE 2 (bases 1 to 7808)  
AUTHORS Tonjes, R.R.  
JOURNAL Direct Submission  
TITLE Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,  
JOURNAL Location/Qualifiers  
1. 7808  
FEATURES  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"  
/tissue\_type="kidney"  
/cell\_line="PK15"  
/gene="gag"  
/protein="gag"  
/codon\_start=1  
/protein\_id="CAA76581.1"  
/db\_xref="GI:3005986"  
/db\_xref="SPTREMBL:O73504"  
/translation="MGQVTTPSLTLDHWTEVRSRAHNLVSQKKGPMQTCASEMP  
TEVGPWSESTSEIILAAKAIIVQGTSSHPDEPYLLTMQDLAEDPPVAKFWLN  
PKRPGPRILALGEEKHSAKEVEPSPIRYEEPTWPEPPVPPVACQVAG  
PSAPGAPVPEAGTASRGATERTDEIILPLRTGPMGPGLOTLQWPESS  
ADLYNKTNNHPEPSEDPRGLTGIVSISMHOPTMDCOLLQTLTTEERRILLES  
RKVPGADRPQIOMETDMGFLTRLDGYNATGARESLKTYRAIYAGLGAARRP  
TNIAKREVMQGNPEPVSFLERLMAFRRTPEPDSQAQSAVALAIGQALDIR  
KTLQLEGLQEAELDLVREAKAYGRTTEBEKQREKEREERDRROENLT  
KTLAAVEGKSSERERDRKTRSGPROSGNLTPLDKQCAVCKEKGHWARNCPK  
KGNKGPKVLALEEDKD"  
/gene="pol"  
/protein="pol"  
/codon\_start=1  
/protein\_id="CAA76582.1"  
/db\_xref="GI:3005987"  
/db\_xref="SPTREMBL:O73505"  
/translation="WGAQISFEQGRREVSVNKPITVLTLDIDERYLSPQKPPDD  
IOSWLEQFOQAAEAGMGLAQVPOVQLKASATPVSRQYPLSRAREIMPVQ  
RLIOGILVPOSPWNTPLLPVKGSTNDYRPVODLREVNKRQVODIHTVPPRYLS  
ALPBERNWTVDLDKAFCLRLHPTSOPLFAFERMDGRTGGLTMTLRDPOGKNS  
PTIPDEALHRLDNFRIOHPOVTLIOYDLDLAGATODCEGKALLILESDGYR  
ASAKKQICREVTYLYGLSLRGQRMLEAKKTYVQIPAPTTAORVEFEGTAFGR  
LMTPGATLAAPLYLTKKEGFSWAPEHOKAFDAIKKALLSAPALADVTKPTLY  
VDEKGVANGVLTQTLGKPRPVAVLSKLDVAGMPCUKAIAVALILVADAKLT  
LQONTIVLHALENIIVRQPDMMNTAMTHYOSLLTERVTFAPALNPATLLPE  
ETPEDEHCHQLELENGVRKLDIDPLTEGYLTMPTDSSYVEEGRMGAAYVQK  
THIWAASSLPEGSAOKAELMALTOALRAEKSINITYDSYARATRAHVHAIYKOR  
GLTTSAGREIKKKEELSLLEALHPKRLAIIHCGHOKAKDLRSQMDRAVAKQA  
AOAVNLLPITETPRAPERQYTLLEWCEIKIDQSETPGTCYTSQKLELPKES  
LELYOUIHRLTLGTHQLOLVRTSPVAVLRPGVADSVKHCVCQVNNPSTIP  
GKRLGSHPGAHMEVDTEVKKPAKYNKYLIVFTDFSGWEAYPTKKETSTVAKKY  
WRKFFRDELYLR"  
/gene="env"  
/protein="env"  
/codon\_start=1  
/protein\_id="CAA76583.1"  
/db\_xref="GI:3005988"  
/db\_xref="SPTREMBL:O73506"  
/translation="MPTLSMRHLPTRGCEPKRLIPLSFASIAFWLTUTTPQASSK  
RLIDSSNPHRLSLTWLIIDPTGVNTSGVAPRGTWPELHFCRLINPAVKSFP

PNLVSYGYFYCCPGTEKERYCGGSESEFRCRMSCVTSNDGKMPISLQDRKFSFVN  
SGPKRYKMYLYKDKSCSPDLDYLYKIFTEKGQENLOKNGSMGIVRYKYGGA  
GSTLIRLRIENGTEPPVAGDVKVLAQGPALPEPNNLPOLITSPITOPPSN  
GTTGILPTNTPNSPVKQKQRLFSIIOCAFOAINTSDPADSCMLCISSEPPY  
BMAKEGKFNVTKEIRNCTQWKSRRKLITLESYSGCTGAPPSHOLCTSTVVEO  
ASENOYLVPQVNRWMACTNGITPCVSSSVFQSKDECVWQIVPVRVYHPEVLEDE  
DYRNRPKREPVSILAVMLIGAVGCTGALITGPOLLEKIGLHAMVEDLR  
ALEESVNLIESITSLSEVVLQNRGLDLPLREGGLCAALKKECCFVDSGAIRDS  
MSKRELERERRERADQWEGEFNFSNPMTTTSLATGPVLVLLLVGCOLIN  
RPAFVRVRVAGVOIMVLRQYCGILSGSERDL"  
BASE COUNT 2077 a 1971 c 2047 g 1713 t  
ORIGIN  
Query Match 96.1%; Score 2839.6; DB 14; Length 7808;  
Best Local Similarity 99.2%; Pred No. 0;  
Matches 2854; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 1 tgccttttaggttagaaccacccttgacagcttgggtctgaacccttaagatgctct 60  
DB 4931 TGGTTTATAGGGTTAGAACACCCCTGACAGATTGGGCTGACCCCTATGAAATGCTCT 4990  
QY 61 aaggggagaccccccccgcttgtagaataatgctctgtacataagtgctgctgtt 120  
DB 4991 AGGGGGACCCCCCGCTGTAGAAATGCTTCTGATAGTGTGATGCTGCTGCTTT 5050  
QY 121 cccagccttgcctcctaagctcaagggctgaagtgtagtgtagaagagctgtagagc 180  
DB 5051 CCCAGCTCTGTTCTTAGGCTCAAGGCGCTGAGGGGTGAGCAACGAGCTGGAAGC 5110  
QY 181 agctccggagagccctactaagagagagagagcttcaagttccacatcgctccaaagt 240  
DB 5111 AGCTCCGGAGGCTTACTAGAGAGAGAGACTTCCAGTTCCATGCTTCCAAAGTGG 5170  
QY 241 gggatcagctcatgttagacgcacgcgtgcaagaaacctcgagctcgttggaaggac 300  
DB 5171 GAGATTCACTCATGTATTAGACCCGACCGTGCAGAAACCTCGAGCTGTTGAGAGGCGC 5230  
QY 301 cctatcgtctacttttgaccacacacagggcgtgaagtgtagaagaaatctccactgga 360  
DB 5231 CTTATCTGCTACTTTTGACACACACAGGCGGTGAAGTCCAAAGAAATCTCCACTGGA 5290  
QY 361 tccatgcattccacgtltaagctgagccacccctccgactcggggtggaagagccgaaga 420  
DB 5291 TCCATGCATCCACAGCTTAAGCTGAGGCGCACCTCCCGACTCGGGGTGAGAGCGCAAGA 5350  
QY 421 ctgagaatcccccttaagcttcgctccatcgctggttccctactactaacaatactccc 480  
DB 5351 CTGAGAATCCCTTAAGCTTCGCCCHCCATCCGCTGTTCTGCTTAACATTAACCTCCC 5410  
QY 481 caggccagtgagaaacgcttataagacgctcgaaccccccatagaacttaccctacc 540  
DB 5411 CAGGCCAGTAGTAACGCTTATTAAGACAGCTGGAACCCCAACGACCTTTATCCCTTACG 5470  
QY 541 tggcgtatattgaccccgatatacggtgtcaactgtaataagcaactcgagtggttgcctc 600  
DB 5471 TGGCGTATTATTGACCCCGATATAGGGGTGTCACTGTAATAGCACTCGAGGTTGCTCCT 5530  
QY 601 agagacacctggtggtcctgaactgcatctcgtcctcgatfgatatacccgctgtttaa 660  
DB 5531 AGAGGACACTGCTGCGCTGAAGCTGCTTCCGCCGATGATTAACCCGCTGTAAA 5590  
QY 661 agcacacctcccaactgctcgttagtgtagtggttctatctgctgcccagagcagagaaa 720  
DB 5591 AGCACACCTCCCAACCTGATGCTGTAATGAGGTTTATTTCTGTGCGCCAGGACAGAGAAA 5650  
QY 721 gagaatactgtggtggtctcgtggaacacctctgtaagagatggaagtcgtacacctcc 780  
DB 5651 GAGAAATACTGTGGGGGTTCTGAGGAATCTTCTGTAGAAATGAGAGCTGCTACACTCC 5710  
QY 781 aacgataagagatggaatgagccgactctctcgaagacgggttaaaatctctctgtgc 840  
DB 5711 AACGATGAGAGACTGGAATGAGCCGATCTCTCCAGAGACCGGTAATAATTTCTCTTGTGC 5770



SOURCE protein.  
ORGANISM porcine endogenous retrovirus.  
VIRUSES; Retrovirus; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.  
REFERENCE  
AUTHORS 1 (bases 1 to 8918)  
Cauderns, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.  
TITLE Establishment and characterization of molecular clones of porcine  
endogenous retroviruses replicating on human cells  
JOURNAL J. Virol. 74 (9), 4028-4038 (2000)  
MEDLINE 20219394  
REFERENCE 2 (bases 1 to 8918)  
AUTHORS Tonjes, R.R.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology,  
Paul Ehrlich-Institut, Paul Ehrlich-Strasse 51-59, Hessen, D-63225  
Langen, GERMANY  
FEATURES  
source Location/Qualifiers  
1..8918  
/organism="porcine endogenous retrovirus"  
/proviral  
/strain="type C"  
/db\_xref="taxon:61673"  
/clone="33"  
/note="grown in PK15 (ATCC No. CCL-33) cell line"  
1..707  
545  
/note="cap site"  
710..727  
/note="g1y4"  
1154..2725  
/gene="gag"  
1154..2725  
/gene="gag"  
/function="retroviral capsid"  
/product="gag protein"  
join(1154..2725,2729..6313,6189..8162)  
/note="viral polypeptide"  
/codon\_start=1  
/protein\_id="CAB65339.1"  
/db\_xref="GI:6688946"  
/translat="MGCTVTPPLSLTLDHMTVEVSRANLTVQVKKGPMTFCASEMP  
TFDVGMPSEGFENSEILAVKAILFQNGSPSHPOEYILTMODLADPPWPMVW  
KPRKPGRIILGCKNKHSAKVEPSPRIPELEPTPEPOPPVPAQAVG  
PSAPPGAVYEGPAGTRSRGATPEPTDELAIPLTYTPMAGGLOPQLQWPS  
ADLNKMTNHPSEDPQRLGLVESLMSHQPWDCDLOLTLTEERELLEBA  
RKNVAGADGPTOLQNEIDMGFPLTRGMDVNTAEGRESLKYQALVAGLRASRP  
TNLAKVEVMGPNPEPVELERLEAFRTPTDPTSEKQASVALAFGOSLDIR  
KTLQREGLQAEALRLDYREAEKYYRRETEKEKREERERBRDRKNT  
KTLAAVEGKSRERERDRPKIRSGPROSGNLGRTPLDQCAVCEKRGHMANCPK  
KGNKGPVLTLEEKDGRGSDPLPEPRVTLKYVGQVPELVDGASHLLDPLKTL  
KEKSWMGATGQKQYFWTRTRIVDLAVGRTHSFVLVPECVPLRLITKGAOI  
SPEQGRVSVNKPITVLTLDDEYRLXSPQVPODIOQWLEOPQMAETKMGAI  
LAKQVPOVIOLKASATPEVSVQYPLSRSEAGIMPVORLIQGLVPOSPWNPPL  
LPVRKPGNDYRPODLREVNKRIVODIHPVNPYNLISALPERRMYVTLIDKDAF  
CLRLHPTSOPLAFEMWRPGTGTGOLMTPLPGFKNPSPTIDEALHRLDRANRICH  
POVTLLOYVDLLLAGATKODICEGTAKALLLESDIGRASAKAOICREPVTLGYS  
LGGGRMLTEARKKTVVOIPAPTRAKOVRELTAGACRIMTPEFATLAPLYLITE  
KGEFSMAHEKQADAIKALLSAPALALPVNTPFTLYDERGVARGVLTQLGFW  
RRPVAVLSKIDPVASGWPCLKAIAVAVALIKADKLTGONITVIAAPHALENTVO  
PPDRMTNARMTHYOSLLTERTFAPALNPATLLPEETDEFTVHDSQLEK  
VRKDLTDLPLTGEVLTWFTDSSVVEKRMAGAAVVDGTHIWASSIPEGSKQAE  
LMALTOALRLAEKSIINYDPRYAFANAHYALHGLTLEISAGREIKRKEILSL  
LEALHLPKRLAIITHCPOKAKDLISRONADVAQOANAVMLPIETPKRDEPR  
KOYLEDQETIKIDQSETEPGTQYISYGEELPHKGBLTVOOILHRLHGLTKHIO  
OLVARTSPYHVLRLPGVADSVVHCPCOLVANASRLPQGRKLGSHGAWEDVTE  
VKPAVGNKYLLVVDPSGWEAVPTKKESTVAVAKLLEIPRPGIPIVISDNG  
PAFVAVOSGLAKIIGIDMKLHCAVRPOSQVGRMNTIETLTKLTETGINDMA  
LPLFVLFVRNTPGOGLPYELLGPGPPVLELAVSASVATLISOPLESKLEMY  
ROBAKOLREASGSGDLOPHRRQVGVSVYRRHAGNLFTBRKGPVLTITPTAV  
RVEGISTVITHTSHVAKLAPPDGSGAEETENPLKRLRLRILPYNNSPGQIHTPLSH  
RHLPTRGEPKRLRILPISFASIAWFLTLITPOASSKRLIDSSNPHRLSLTWLIDP

DTGTVNSTRGVAPRGITWPELHFLRLINPAVKSTPPNLYRSYGYCCPTEKEXYC  
GDGSEFCRMSQVTSNDGDKWPIISLDQKFSFVNSGPKYKWKILKDKSCPSD  
LDYKISFTEKQENIKWIMGMSGIVFYKGGAGSTLIRLITGTEPPYAVG  
PKVLAEGOPALPEPPHNLVPOLTSILDTIOSPNSGTGLIPNTRNSGPVKT  
GORLSILQGAFOALINSDPDATSCWCLISGPPYEGMAKEKRENOCTW  
GSRNKLITLVSGKTCIGKAPPSHOHCISYVYEQASENOYLVPGYRRMACNTGL  
TPCVSTSVFNSKDRCVWQIVPVRYTHPEEVYDEYTORINRKRREVSLTAMLS  
LGTAVGCTGATLITGPQOLEKGGELHAMTEDLRLAESVSNLESLTSLSVL  
ONRGLDILFTEGLCAALKKECCFYVDHSGAIDMSKILRELRERRERADOG  
FEGFNRSFWMNTLLSALGPLVLLLLITGPCLINRFVAFRRVSVAQOIMVLRQO  
YGLLSQGETDL"  
2729..6313  
/gene="pol"  
/function="retroviral protease and polymerase"  
/product="pol protein"  
2729..6313  
/gene="pol"  
6189..8162  
/gene="env"  
/function="retroviral envelope"  
/product="Env protein"  
6189..8162  
/gene="env"  
8212..8918  
/gene="gag"  
8811..8816  
polYA\_signal  
BASE COUNT 2327 a 2245 c 2334 g 2012 t  
ORIGIN

Query Match 95.1%; Score 2811.6; DB 14; Length 8918;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 2914; Conservative 0; Mismatches 24; Indels 78; Gaps 1;

QY 1 tgccttttaggttagaaccacccttgagagacttggtgcgcacccctatgaattgctc 60  
DB 5826 TGCCTTTTAGGCTTAGAACAACCCCTGACAGATTGGCGTGACCCCTATGATTGCTCT 5885  
QY 61 acgaggagaccacccttgtagaacttgctctgtacatagtgctgtagtgcctt 120  
DB 5886 ACGGAGGACCCCGCGCTGTAAGAAATTCCTCTGACAAAGTGATGCTGCTCTT 5945  
QY 121 cccagcttgctctcctagagctcaagagcctgaagtgagagcaagagcgtggaagc 180  
DB 5946 CCCAGCCTCTCTCTCTGAGGCTCAAGGCGCTGAGTGAGGCAACGAGCTGGAAGC 6005  
QY 181 agctccgagagagcctcctcagagagagagagactgcaagttccacatcgcttcaagt 240  
DB 6006 AGCTCCGAGAGCCTACTCAGGAGAGAGAAACTTGCAGTTCCACATCGCTTCAAGTG 6065  
QY 241 gagattcgtctatgtagacgcacccgtagcaagaaactcgagactcgtagaagagac 300  
DB 6066 GAGATTGAGTCTATGAGACGCCACCGTCAGAGAACTCGAGACTCGTGGAAGGCGC 6125  
QY 301 ctatctcgtactcttgacacacacacagcgtgtggaagtcggaagaaatctccactgga 360  
DB 6126 CTTATCTCGTAACCTTGAACACACAGGCTGTGAAGTGGAAGGATCTCCACCTGGA 6185  
QY 361 tccatgcatccacgtagaagcttgagcgcacactcccgagactcgaggtggaagacgaaaga 420  
DB 6186 TCATATACATCCCACTTAAGCTGGCGCACCTCCGACTGGGGTGAGAGCGGAANAAGA 6245  
QY 421 ctggaatccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 480  
DB 6246 CTGGAATCCCTTAAGCTTGCCTCCATCGCTGCTCTCTTACTTAAACAAATACCTCC 6305  
QY 481 cagagcagtagtaagcgtctatagagaagtcggaaccccatagaaccttaccctcacc 540  
DB 6306 CAGGCGATAGTAACGCTTATAGACAGCTGGAACCCCATAGACCTTATACCTTACC 6365  
QY 541 tgcctgattatagacccctgatacagagtgctcactgtaaatagcactcgaggtgctcct 600  
DB 6366 TGGCTGATTATTGACCTCGATAGCGGGTGACCTTAATAATAGCACTCGAGGTGCTGCT 6425

QY	601	aagagcaactcgtgmggctcgaactgatacttcctcgcgcgcgattgataaccccgctgttaa	660
Db	6426	AGAGGACCTGGGGGCTCGAACTGCAATTTCTGCCGATTGATTAACCCCGCTGTTAA	648
QY	661	agacacacccccaaccttagccgtgtagtatagtgtctctctcgtcgtccagcaagagaa	720
Db	6486	AGCACAACCTCCAACTATCTCGTAGTTATGGTTCTATTCTCTCCAGGCACAGAGAA	654
QY	721	gagaatactgtgagggtctcgtgggaatcctctcgtlagagatggaatcgtcactcc	780
Db	6546	GACAAATACTGTGGGGATTCTGGGGAATCTTCTGTAGAGATGCACTGCTCACCTCC	660
QY	781	aacgatgagagatctggaataggccggtctctccaggaaccggtgaaattcctctgtc	840
Db	6606	AACGATGAGACTGGAATAAGGCGCATCTCTCCAGGACCGGGTAAATTTCTCTTTGTC	666
QY	841	aattccggcccggggaagatbacaaatgataaacatataaaga taagatcgtctccca	900
Db	6666	AATTCGGCGCCGGGCAAGTACAAACTGATGAAACTATATTAAGATTAAGACTCTCCCA	6722
QY	901	tcaagctatgattactaaagataaagttcacctgaaaaagaaacagaaaaatltcaa	960
Db	6726	TCAGACTTATGATTAATCTAAAGATTAAGTTTACTGATAAAAGAAACAGAAATATTCA	6788
QY	961	aagtgataaagatgatatgctcgtgggaaatgtttttaaataatgctggggagcagg	1022
Db	6786	AAGTGATTAATGATATAGCTGGGGGAATTAATTTTATTAATATGCGGGGAGCAGGG	6848
QY	1021	tcaacttaaacatccgcttaga taagaacgggagcagaaccctctggaataggga	1080
Db	6846	TCACATTATACATTGCGCTTAGATATAGATAGAGCGGGGCAAAACCCCTGTGGAGTGGGA	6907
QY	1081	cccgaataagtactggtcgtgaacaggggcccccgcccttgtagccacgcataactgcg	1140
Db	6906	CCCGATTAAGTACTGGCTGAACAAGGGGCCCCCGGCTCTGAGGCACCGCATTAATTGCC	6966
QY	1141	gtgcccgaatbaactgcgtctcggtcgtgacataacacagccgctagcaacagtaacct	1200
Db	6966	GTGCCCCAATTATTAACCTCGCTCGGCTGTGCATTAACAGATGCGCTAGCAACGGTATACACT	7022
QY	1201	ggattgatctctacaacacgcctagaaatccccagtgctctcgtttaagacaagag	1260
Db	7026	GGATTGATTCCACCAACAGCCCTTAGAAACTCCCGAGGTGTCTCTTTAAGACAGGACAG	7088
QY	1261	agaactctcagctctcatccaggagagcttccaaagcatcaactccacgcgacctatgpc	1320
Db	7086	AGACCTCTCAGTCTCATTCAGAGGAGCTTTCCAAAGCATCAATCCACGACCCCTGATGCG	7145
QY	1321	acttttctgttgcttctctactccagggcctccctatatagaaggagatgctaa	1380
Db	7146	ACTTTCTTGTGTGGCTTTGCTATCTCCAGGGGCTCTTATATATAGGGGATGGCTTAA	7205
QY	1381	gaaagaanaatcaatgtagaccaagaagcatagaaatcaatgtacatgggggtcccgaaat	1440
Db	7206	GAAAGAAATTCATGTGACCAAAAGACATAGAAATCAATATGTACATGGGGGTCCGCAAT	7265
QY	1441	aagcttaacccctacgtgaagttcccggaagggagacatgataggaaaagctccccatcc	1500
Db	7266	AAGCTTACCTCCTACAGTAAGTTTCGGGAAAGGAGCATGTGATGAGAAAAGCTCCCCATCC	7325
QY	1501	caaccaacactctgactatagtaactggtttatagacagggcctcagaanaatcagatatta	1560
Db	7326	CACCAACACTTTTCTATAGTACTGTGGTTTATAGACAGAGCTCCAGAAATATCAGATTTTA	7385
QY	1561	gtacctgttataacaggtgtagatgataatcagggtttaacccctcgtgtttccacc	1620
Db	7386	GTACCTGTTATTAACAGGTGGTGGGCAATGCAATATCTGGGTTAAACCCCTGTGTTCCACC	7445
QY	1621	tcaagtctcaaccacatccaaagaattgtgtctatgltccaatcgtccccgaatgtac	1680
Db	7446	TCAGTCTTCACCAATCCAAAGATTTCTGTGTATGTGTCCTCAAAATGCTGCCCGAGTGTAC	7505
QY	1681	taccatctcgaagaatggtcctctgabaatalgacatcgtgataaccagaccaaaga	1740

[illegible]

```

Db      8586 GAGGCGGCGCTCGATATTTAAATGATTGGTCCACGAGCGCGCTCGATATTT 8645
Oy      2760 -----ggagcgagcgctcgatattttaaattgattggaag 2802
Db      8646 TAAATGATTGTCACAGGAGCGGCTCGATATTTAAATGATTGTTGACG 8705
Oy      2803 cacaagcttctgtgaacccataaagctgcacgattccgactcgagcgagcagc 2862
Db      8706 CACAGGCTTGTGTGACACCCCAATAAGCTGTCGCGATTCCGCACTCGGCGGAGTC 8765
Oy      2863 cctcaccctcgctggtgtagacactgtagcccgagcgcttgataataaactctt 2922
Db      8766 CTCTACCCCTCGCTGTGTCAGACTGTGGCCCGACGCGCTTGAATTAATAATCCTCTT 8825
Oy      2923 gctgttgatcaaaa 2938
Db      8826 GCTGTGTGCAACAGA 8841

RESULT 9
LOCUS   PEN133817 8849 bp DNA VRL 07-JAN-2000
DEFINITION Porcine endogenous retrovirus type C proviral gag, pol and env genes and LTR (class A, clone 42).
ACCESSION AJ133817
VERSION AJ133817.1 GI:6688947
KEYWORDS env gene; env protein; gag gene; gag protein; pol gene; pol protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
REFERENCE 1 (bases 1 to 8849)
AUTHORS Czauderna, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and Toenjes, R.
TITLE Molecular Characterization of Human-tropic and Replication-competent Porcine Endogenous Retroviruses
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8849)
AUTHORS Toenjes, R.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Toenjes R. R., Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY
FEATURES
     source          1..8849
                        Location/Qualifiers
                        1..8849
                        /organism="porcine endogenous retrovirus"
                        /strain="type C"
                        /db_xref="taxon:61673"
                        /clone="42"
                        /note="grown in PK15 (ATCC No. CCL-33) cell line"
                        1..668
                        /note="cap site"
                        671..688
                        /note="gly4"
                        1115..2686
                        /gene="gag"
                        1115..2686
                        /gene="gag"
                        /function="retroviral capsid"
                        join(1115..2686,2690..6271,6150..8132)
                        /product="gag protein"
                        /note="viral polypeptide"
                        /codon_start=1
                        /protein_id="CAB65340.1"
                        /db_xref="GI:6688948"
                        /translation="MGQVTPPLSLTLDHMFVRSRAHNLISVQKKGMOTFCASEMP
                        /EDVGRPSGRTNSSEILAVKALITOTGSGSPDEPPIITLMDIAEDPPPVKRMNL
                        /KPKRGPRTIALGKRNKSHAEKVEPSRPIYIEPPTWPEPQVPPPPVPGQAVRG
                        /PSAPGAPVVEGPAAGTRSRGATPERKDEIALILRLTYGPMGGOLOQVWPFSS
                        /ADLYNWKTNHPFSEDPQRLTGLVESLMSFHSQPTWDDCQQLQTLFTTERERILLEA"
CDS
     misc_feature    671..688
                        primer_bind
     gene            1115..2686
     mat_peptide     1115..2686

```

```

RKNVPGADGREPTOLONEIDMGFPILTRPGMDYNTAEGRSESLIKYRQALVAGLRASRP
TNLAKEVMQGNPEPSPVLEERLMEAREPTPEPTSEAKOASVALTAGSLODR
KTLRLGLOAEALDLYREAEKYVRREREKEERERERERERERERERERERERER
KTLAAVBEKSSRRERERERERERERERERERERERERERERERERERERERER
KGNNGPVALLAEEDKDRSDPLPEPRVLTAKGQVEEVLVDGASHVLDPLGUL
KEKSWMGAEDGQRPWTKRTVRRVDTAGVGRTHSEFLVLPBPVLLGRDLTKGAGI
SFEQGRSEVSNNKPIVTLVDLDEVRLESPQVKPQDQIOMLEQPOQMAEPTAG
LAKOVPOVLOLKASATPVSVROYLSREAREGIMPVORLOOGLILVPOSPNTPL
LPVRKPNNDYRPVODI.REVKNRYVODIHPVNPVNYNLISALPERMRYTLDLDAF
CLRLHPTTSQPLPAFERMDPGTGRGQTLTRRLPDGFENSTPIDEALHRLANFKIO
POVTLVYVDLVLGATKQDCLEGTALLLESDLYRASAKAQICREBYVLGYS
LRGQRMVTEARKKTVVOIPAPTTAKGVREFGTAGCRIMIPQFALAPLPLTKE
KGESMABEHQKADAIKALISALPDVETPFLLYDERGAVAGVLTOTLGM
RPPVAYSLKLDPAVSGMPYCKAIIAVALIKVADKLTGLONTVTAHALEIVVO
PDRWMTNARKTHYOSLITERVFAAPALNPATLLPEFDEPVDHCHALLIEEG
VRKDLTDLPLGELVLTWFTDSSVYVGGKRMAGAAVYDGRTRTASSLPQTSQAK
LMALTOALRLAEKSTINTYDSRIAPATVHGAITYKQGLTSAGEIKKEISL
LEALHPKRLAIHICPGHOKAKDLISRGNOADVAQAQAVNLPLIETPRAPER
ROYTLEDMOEIKIDQFSETEPGCTSYGKELIPHEGLEYVOQIHLRTHGKHQ
OLVETSPYHVLRLPGVADSVYHCVPCOLVANSRILPPGRLRSGHMEVDFTF
VKPAKYNKYLIVFDNESGWEAVYPRKKESTVYAKKILIEIFPRGCIKVISDNG
PAPVAYOSGIKTLGTDMLKHCARPOSSQOVRNMRRTKETLTKLTGTTGINDMA
LPLVLEVRNTPQGFGLPYELLVGGPPLVEIASVSHADVLISQLFSLKALEW
RORAMKOLREAVYSGEDLOVPHRQVDSVYRRHAGNLETWKGYLVLTPPTAV
KVERIPTWISHAVKPAAPPDSGKAKETEMPKLRLHRYVYVSNNSSMHTLSRR
HLPIRGKPKRLKLPISFASIAFWLTLSIPVOYNGKRLVDSNPKSLTWLITDGG
TGINSTOGAPLPGTWMPPELVYGLRSVITGINDOATPPVLAHYGYVCGPNNNE
YCGNPDPFCKOMSCITSNQNMKWPISQODRYSYVNNPTSNQNYHGRKMDQ
ORVOKDVNRKQISCHSIDLDTLKTIFTEKQOEIYQWVGISGLVYGGSGKKIS
VLTIRLREIOMEPVPAIPNPKGLADEGPPLOEDRPNEDVNTTSGVSETKNKS
KTGAKSLILOAGAOALNSTPTEATSCWCLASGPYVYGMARGFENYKEDDC
TWGSONKSLITLEVSGKGTIGMPPSHOHCNHEANRTSESOYLAVGDRMKWAC
GLTPCVSTIVENOTKDCVMOIVPVRVYPERKAVLDKYRYRNRAREPISLTIAV
LGISVANGVGTTRALITGTVOQLKEGSLNHRITVEDLOALEKSVNDESLISLV
VLQNRKGLDLFEKGDVCAKKECCFYVDHSATIGDSKRLERLERREAEAO
GMFGEWENRSFMTWLTLLSALPLPLVLLLTIVGPDILNRFVAFRBRVSAVOIMVLR
QOYGLSGEFTDL"
2690..6271
/gene="pol"
/function="retroviral protease and polymerase"
/product="pol protein"
2690..6271
/gene="pol"
/gene="env"
6150..8129
/function="retroviral envelope"
/product="Env protein"
6150..8129
/gene="env"
8182..8849
LTR
polyA_signal
BASE COUNT 2355 a 2281 g 2022 t
ORIGIN
Query Match 67.8% Score 2004.4 DB 14; Length 8849;
Best Local Similarity 81.6% Pred. No. 0;
Matches 2480; Conservative 0; Mismatches 406; Indels 152; Gaps 8;
Oy      1 tgccttttaggttagaacacccctgacagcttgaggctgacccctatgaattgctc 60
Db      5787 TGCCTTTTAGAGGTGAGGAACACCCCTGCAGATTGGGCTGACCCCATGAAATGCTCT 5846
Oy      61 acggggagaccccccccgctgtgtagaattgctctgtcatatgagtgtgctctt 120
Db      5847 ACGGGGGAGACCCCCCGTGTGAGAAATTTGCTTGTCATAGTATGCTATGTGCTGCTT 5906
Oy      121 cccagcccttgctctcagagcgctcagatgaggtgaggaagcagcagcgtggaagc 180
Db      5907 CCCAGCCTTTGTTCTCTAGGCTCAAGCGCCTCGAGTGGGCTAGAGCAACGACGTTGGAAGC 5966
Oy      181 agctcggagagcctactcagagagagagacttgcaagtccacatcgctccaagt 240

```



Db 5967 AGCTCCGGAGGCTTACTACAGAGAGAGACTTCAGATTCCACATCGCTTCCAGTTG 6026  
 Qy 241 gagatcagctctatgttgaagcgaagtcgaagaacctgagaatcgttgaagggac 300  
 Db 6027 GAGATTTCAGTATATGTGTAGCCGCCACCGTGCAGGAAACCTGAGACTGGTGAAGGCC 6086  
 Qy 301 ctatcctgtaattttgacacacgaagcgtgtgaagtcgaagaaatcccaacctgga 360  
 Db 6087 CTTATCTGTACTTTTGAACACACACAGGCTGTGAAGTGAAGAAATCCCACTGGA 6146  
 Qy 361 tccatgcatcccaagcttaagcttgcgcgaacctcccgactcgggttgaagcgaaga 420  
 Db 6147 TCCATGACATCCACGTTAAGCGGGCCACCTCCGATTTGGGGTGGAAACCCAGAAAGA 6206  
 Qy 421 ctgagaatcccttaagcttgcgcgaagtcgcgttgccttaacttaacaataactccc 480  
 Db 6207 CTGAATAATCCCTTAAAGCTTGCCTCCATCCGCTGCTCTTACTCTGTCAATTAACCT 6286  
 Qy 481 caggccagtaagtaaacgccttatagacagctcgaaccccaataaccttaacctacc 540  
 Db 6287 CAAGTTAATGTAAGAGCTTGTGACAGCCGAACTCCATAACCTTAATCTCAAC 6326  
 Qy 541 tggctgatatgaacctgatacgggtgtacactgataaataagcagctgagtgctccct 600  
 Db 6327 TGGTACTTACTGACTCGGATACAGGATTAATTAACAGACTCAAGGAGGAGCTCC 6386  
 Qy 601 aagagcaccgtgtgcctgaacgtcaattcgtcctcgatgtalttaaccccgctgta- 658  
 Db 6387 TTGGGAGACTGTGCGCTGGAATTAATGTCTGCTGCTGATGATTAATCCGCTCAT 6446  
 Qy 659 ----aaagcacacctcccaactagtcgtagtatagtgtcttctgctccgaag----c 711  
 Db 6447 GACCAAGCCACACCCCGATGATCTCGTCTTACCGGGTTTACGTTTGGCCAGGACC 6506  
 Qy 712 acagagaagaagaataactgtgtggtgtcctggaagacccctctgtaagagatgagctgc 771  
 Db 6507 CCAATATATGAAGATATGTGGAATCTCAGATTTCTTTTCCAGCAATGAGCTGC 6566  
 Qy 772 gtcaacctcaagatgagactgtgaatgagcgaatcctcctcgaagcgggtlaaattc 831  
 Db 6567 ATAACTTCTAATGATGGGAATTTGGAATTTGGCACTCTCAGCAAGAGATTAAGTTAC 6626  
 Qy 832 tcccttgcaattcc----- 846  
 Db 6627 TCTTTTGTTAACAATCTCACCAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 6686  
 Qy 847 ----ggccgggcaagtaacaaatgatatgaataataataaagatgagcttcccatc 902  
 Db 6687 GATTGGCAACGCGGTACAAAAGATG-TACGAATTAACCAATTAAGCTGTCAATTCGT 6745  
 Qy 903 agacttgaattatctaaagataagttcactgaaagaaagaaagaaagaaatatacaaaa 962  
 Db 6746 AGACTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6805  
 Qy 963 gtgataaagtatgtagctgtggaatagtlttita----ataatggtggtggtggtggtg 1019  
 Db 6806 GTGGGTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1079  
 Qy 1020 gtccactttaaactatgctgtggtatgagagagcgggagacacccctctggaatggtg 1079  
 Db 6866 ATCTGTTCTGACTATGCGCTCAGAAATGAAGATCAAGATGGAACCTCGGTTGCTATAG 6925  
 Qy 1080 acccgataaagtagctgtgtgaagagggcccggtggtgagcgaacccgataactgtcc 1139  
 Db 6926 ACCAAATTAAGGTTTGGCCCAACAGGACCTCCATCAATCAAGA----- 6967  
 Qy 1140 ggtgcccataaacctcgtgtggtgctgacataacacgctcgtgagacagctaacac 1199  
 Db 6968 -----ACAGAGGCCATCTCTTAACCTCTGATTAACATTAACATCAACCTC 7009  
 Qy 1200 tggatgtatcttaccacacagcctagaagaatcccgagtgctcgttgaagcagagaca 1259  
 Db 7010 TGGATGATGCTCCCACTGA-----GCCATACTACATTAATTAATTAACAGGGCC 7054

Qy 1260 gagacttcaagtcataccaggaagcttcccaagcatcaactccagcagccctgatagc 1319  
 Db 7055 GAAACTTTTATAGCTCTATCCAGGAGGCTTTTCAAGCTCTTAATCTCCAGCATCCAGAGG 7114  
 Qy 1320 caacttcttctgt 1379  
 Db 7115 TACCTCTTCTGT 7174  
 Qy 1380 agaaagaanaatcaatgtagcacaagagcaataaatacaatgtagcagagagagagagag 1439  
 Db 7175 AGGAGGGAATTTCAATGTGCAAGAGACATTAAGACCAATGTACATGGGAGATCCCAAAA 7234  
 Qy 1440 taagcttaacctcaatgaagtttccgggaagggagacatgataagaaaagctcccccac 1499  
 Db 7235 TAAGCTTACCTTACTGAGGATTTCTGAAAAGGACCTGATGAGATGTTCCCCATC 7294  
 Qy 1500 ccaaccaacacctgtgtatgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1559  
 Db 7295 CCACCAACACCTTTGTAACCCACACTGAAGCCTTTAATCAACCTTGAGAGTCAATATCT 7354  
 Qy 1560 agtaacctgataaacaagtggtgtggtgcatgatactgtgtgtgtgtgtgtgtgtgtgt 1619  
 Db 7355 GGTACCTGGTATGACAGAGGTGGGCTATTAATTAATTAATTAATTAATTAATTAATTA 7414  
 Qy 1620 ctaagtttcaacaaatccaaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1679  
 Db 7415 CTGTGTTTCAACCAATTAAGATCTTTGCGTTATGTTCAATTTGTCCTCCCGGGGTGA 7474  
 Qy 1680 ctacactcctggaagagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1739  
 Db 7475 CTACTATCCCGAAAAGAGAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTA 7534  
 Qy 1740 agaacccgatacccttaacactgataagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1799  
 Db 7535 AGAGCCCATATCTCTACATCACTAGCTATTAATTAATTAATTAATTAATTAATTAATTA 7594  
 Qy 1800 aacaggaagctgtccctgtatcaagagcccaagcagcagcagcagcagcagcagcagcag 1859  
 Db 7595 AACAGGAAGGCTGCGCTTAATCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7654  
 Qy 1860 acatggtgcatagacgaagagatcctcgagccttaagagagctgtgtgtgtgtgtgtgt 1919  
 Db 7655 ACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7714  
 Qy 1920 gtccctgacttctgt 1979  
 Db 7715 ATCCCTAACCTCTCTATCTGAAGTGTCTACAGAACAGAGGGGTTATGATCTTTAT 7774  
 Qy 1980 tctaagagaagtggt 2039  
 Db 7775 TCTAAAGGAAGAGGAGGT 7834  
 Qy 2040 ctcaagagccatcagagactcctgaacaaagcttaagaaagcttaagagagctgtgtgtgt 2099  
 Db 7835 CTGAGGAGCCATCAGAGACTCTCATGACAAAGCTTTAGAGAAAGGTTAAGAGGCGTCGAG 7894  
 Qy 2100 ggaagaagagctgtgacaaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2159  
 Db 7895 GGAAGAGAGGCTGACAGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7954  
 Qy 2160 caacctgttctgt 2219  
 Db 7955 CACCTGTCTTGTCTGTGACAGGAGGCTCTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTG 8014  
 Qy 2220 gcttgtcttaataataggt 2279  
 Db 8015 GCTTGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8074  
 Qy 2280 catgttacttaagcaacagtagcaagccttctgtgcaagagaaagaaagaaagaaagaaag 2339  
 Db 8075 CATGTACTTAAGGACACAGTACCAAGGCTTCTGTAGCCAAAGAGAAACTGACCTTAAGCC 8134



QY 2340 ttccagcttaagatagaaactaaacagaagaagtgagggaatgaagatgaagaa 2399  
DB 8135 TTCCAGTTTAAAGATAGAACTATTACAGACAGAAAGTGGGAATGAAGAGATGAAA 8194  
QY 2400 atgcaactaaccccccagaagaccagaagaatataaagctctaaatgccccgaat 2459  
DB 8195 ATGCAACTAACCCCTCCAGAACCCAGGAATTAATAAAGCTTAATGCCCCCAAT 8254  
QY 2460 tacaacccctgctgctccagtaaatagtagaaggtcacactctatgttccagg 2519  
DB 8255 TCCAGACCTGCTGGCTCCAGTAATAGTAAGAAAGTCCACTTCTATTTCCAGG 8314  
QY 2520 cctgctatccctgagctaaagtaagaataaagaaatgagtaactatogcttaactgat 2579  
DB 8315 CCTGCTATCTCGGCTTAAGTAAGTAACAGGAATGAGTTGACTTAATCGCTATCTGAT 8374  
QY 2580 tctgtaaacactgactgagacatagaagaattgataacattgacagcccttagtgacct 2639  
DB 8375 TCTGTAACACTGACTGACCACTAGAAAGTAATGATACACTGACAGCCCTAGTGACCT 8434  
QY 2640 atctaacctgcaatctgcaactctgcccagaagcccagacgacatgagacctccgagac- 2698  
DB 8435 ATCTCACTGCATCTGCTGCACTTGCCAGAGGCCAGCGAGTCCGAGACTCCGAGACT 8494  
QY 2699 -----tatttaaatgattggtccac 2720  
DB 8495 ATTTTAAATGATTGCTCACGAGCGCGGCTCTCGATATTTTAAATGATTGCTCAC 8554  
QY 2721 ggaagcgaggctcgcgaatttaaatgattggtccacgagcgagcggtctcgaatc 2780  
DB 8555 GGAAGCGGCTCTCGATATTTTAAATGATTGCTCACGAGCGGCTCTCGATATT 8614  
QY 2781 ttaaatgattggttctgacgacagcggttctgtaacccataaagctgtccga 2840  
DB 8615 TTAATGATTGTTGTTGACGACAGCGCTTGTGTAAGCCCATTAAGCTTCCCGA 8674  
QY 2841 ttccgcaactgagcgagcgaactctcctccctgctggtgtaagactgtgagcccgag 2900  
DB 8675 TTCCGCACTCGGCGCCAGCTCTCTACCCCTCGTGTGACGACTGTGCGCCCGAGCG 8734  
QY 2901 cgcttgataaaatcctctgctgttgcatacaaaa 2938  
DB 8735 CGCTTGAATATAAATCTCTGCTGTTTGATCATAGA 8772

RESULT 10  
A66559  
LOCUS A66559 1974 bp DNA PAT 29-MAR-1999  
DEFINITION Sequence 9 from Patent WO9740167.  
ACCESSION A66559  
VERSION A66559.1 GI:4538112  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unclassified.

REFERENCE  
1 (bases 1 to 1974)  
AUTHORS Galbraith,D.N., Hamworth,C., Lees,G.M. and Smith,K.T.  
TITLE PORCINE RETROVIRUS  
JOURNAL Patent: WO 9740167-A 9 30-OCT-1997;  
Q ONE BIOTECH LTD (GB)  
FEATURES  
Location/Qualifiers  
Source 1..1974  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 534 a 489 c 483 g 468 t  
ORIGIN

Query Match 66.7%; Score 1970.8; DB 6; Length 1974;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgcatcccaagcttaagctgagccacctccgactcgagggtgagagagccgaaagactg 423

DB 1 ATGATCCCAAGTTAAAGCTGGCGCACTCCGCACTGGGGTGGAGAGCGGAACAGT 60  
QY 424 agaatcccttaagcttgccctcaatccgctggttcccttaacttaacaactcccg 483  
DB 61 AGAATCCCTTAAGCTTGCGCTCCATCGCTCGTCTCTTAATCAATTAATCTCCAG 120  
QY 484 gccagtagtaaacgcttataagacgctgaaccccatagaccttataccctactcg 543  
DB 121 GCCAGTGTAAAGCGCTTATAGACAGCTCCAAACCCCATAGACCTTATCCCTACTCG 180  
QY 544 ctgattatgacccctgatacgggtgtcactgtlaaalagacactgaggtgtgctccaga 603  
DB 181 CTGATTATTACCCCTGATACGAGGTGACACTGTAATAGCACTGAGAGTGTGCTCTAGA 240  
QY 604 ggcacccgtggtgacctgaactgcattctgctccctgatttaaccccgctgttaaagc 663  
DB 241 GGCACCTGGTGGCTTAACCTGCAATTTCTGCTCCGATGATTAACCCCTGTTAAAGC 300  
QY 664 acacccccaactagtcgtaagttatggttctatctgctgcccagagcaagaaagag 723  
DB 301 ACACCTCCCAACTAGTCCCTAGTTATGGTTCTATTGCTGCCAGGACAGAGAAAG 360  
QY 724 aaatcctgaggggttctgaggaatccctctgtaagagatgagctgctcaactccaa 783  
DB 361 AAATAGTGTGGGCTTCTGGGAATCTCTGTAGAGATGAGAGTGGCTCACCTCCAA 420  
QY 784 gatgagagctggaatgagccgactctctccagagccgggttaaattcccttgcatt 843  
DB 421 GATGAGACTGGAAATGGCGATCTCTCCAGAGCGGGTAAATTTCTTGTCAAT 480  
QY 844 tccgcccgggcaagtagaacaatgtagaacataaagaatgaagctgtccccaata 903  
DB 481 TCCGGCCCGGCAAGTAGAACAAATGATGAACCTATTAAGATAGAGCTGCTCCCATTA 540  
QY 904 gacttgattatcctaaagatgattcactgaaagagaaacagaaatattcaaaag 963  
DB 541 GACTTGAATATCTAAAGATTAAGTTCACTGAAAAAGCAAGAAATATTCAAG 600  
QY 964 tggataatgtagtagcgtggggaatgattttaaataatgagggggaagaaggttc 1023  
DB 601 TGGATTAATGATGAGTGGGGAATGTTTTTATTAATATGAGCGGGACAGGGTTC 660  
QY 1024 accttaacatcgccttgatagtagagcgggagcaaaccccgctgacatggaaccc 1083  
DB 661 ACTTTAACCATTCGCTTAGATAGAGAGGGGACGAAACCCCTGTGCAATGGAGCC 720  
QY 1084 gataaagtaactgctgaaacaggggcccggccctgagagcaacgataactgtccgtg 1143  
DB 721 GATTAAGTACTGCTGAACAGGGGCCCGGCTTGAGCCACACCGCATTAATTCGCGTG 780  
QY 1144 ccccaatlaacctcgtctgagctgacataacacagcgccctagcaacagatccactgga 1203  
DB 781 CCCCAATTAACCTCGCTGGGCTTGACATTAACACACCCGCTTAGCAACGTCCACTGGA 840  
QY 1204 ttgattccacaacacgactagaactcccaagtggttcctggttaagaacagagaaga 1263  
DB 841 TTGATTCTTACCAACAGCGCTAGAAACTCCAGGTTCTCTGTTAAGCAGAGAGAGA 900  
QY 1264 ctcttgaactcactcaagggagcttccaagccatcaactccacgacctgtagtccact 1323  
DB 901 CTTCTAGTCTCATCCAGGAGCTTCCAAAGCATCACTCCACGACCTGATCCCACT 960  
QY 1324 tctctctgttgcttcttaccctccagggctcctcttataatgagggagatgctaaaga 1383  
DB 961 TCTTCTGTGGCTTTGCTATCTCAGGCGCTCTTATTAAGAGGAGTGGCTTAAAGAA 1020  
QY 1384 agaaaaatcaatgtagcaacaagaagatagaatgaatgtagatgaggggtccgaaataag 1443  
DB 1021 AGAAAAATTAAATGTGACCAAAAGAGATAGAAATCAATGTACATGGGGGTCCGAAATTAAG 1080  
QY 1444 cttaacctactgaagtttccgggaaggagatgcatagagaagctcccccataccac 1503



QY 844 tccgccccggaagtaacaaatgatbnaactatataagaagagctgtcccatca 903  
 Db 481 tccgccccggaagtaacaaatgatbnaactatataagaagagctgtcccatca 540  
 QY 904 gaactagattactaaagaatgttcaatgaaagaagaaacaggaataatltcaaaag 963  
 Db 541 GACTTAATATATCAAAAGATAGTTTCACCTG--AAAGGAAACAGGAAATATTTCAAAAG 598  
 QY 964 tggataatgtaagatgagctgggaatagtttttaataatgagcgaggagaggttc 1023  
 Db 599 TGGATTAATGATGATGAGCTGGGGAATAGTTTTTT-ATTATATGCGGGGAGGAGGCTC 657  
 QY 1024 acctaaacctgccttaagatagaaacgggagcaaacacccctgtgcaatggagacc 1083  
 Db 658 ACTTTAACATTCGCCCTTAGATAGAGACGGGGACAAACCCCTGTGGCAATGGGACCC 717  
 QY 1084 gataaagtaactgctgaacagggccccggccccctggagccacgcataactgtccggtg 1143  
 Db 718 GATTAAGTACTGGCTGAACAGGGGCCCGCCCTGGAGCCACCGCATTAATTTGCGGCTG 777  
 QY 1144 ccccaatlaactgcgtgcggcctgacataacacagcgcctagaaacagttaccactgga 1203  
 Db 778 CCCCAATTAACCTCGCTGCGGCTGACATTAACACGCGCTAGCAACAGTACCACTGGA 837  
 QY 1204 ttgattcctlaacacacgacctagaacactccacagtggttcctgtlaagacagagaga 1263  
 Db 838 TTGATTTCTACCAACACGCGCTAGAAACTCCCAAGGTGTCTCTGTAAAGACAGGACAGA 897  
 QY 1264 ctcttcagctcatcagaaggagcttccaaacatcaactccacgcacccctgagccact 1323  
 Db 898 CTCTTCAGTCTCATCCAGGAGCTTTTCAAGCCATCAATCCACCGACCTGTATGCCACT 957  
 QY 1324 tctctgtgttgctgttgcatacctcaaggcctccttatatagagggatgagttaaagaa 1383  
 Db 958 TCTTCTGTGTGGCTTTGTGCTATCTCTCAAGGCTCTCTTATTATGAGGGATGGCTAAAGAA 1017  
 QY 1384 agaaatcaatgtgacaaagaagacatagaaatcaatgtaacatgagggggtcccgaaataag 1443  
 Db 1018 AGAAATTTCAATGTGACCAAAAGACATAGAAATCAATGTACATGTACATGTGCGGAGTAAAG 1077  
 QY 1444 cttaacctacatgaagttccgggaaggagacatgacatagaaagagctcccccacccac 1503  
 Db 1078 CTATCCCTCACTGAAGTTTCCGGGAGGAGGACATGATAGGAAAGAGCTCCCATCCAC 1137  
 QY 1504 caaacacttgcatactagctgtgttataagagagccctcagaagaatcagattatagta 1563  
 Db 1138 CAACACCTTGTGCTATGATGCTGTGTTATGAGCAGGCTCAGAAATWCAATTTATGTA 1197  
 QY 1564 cctgttataacagtggtggaatgacatgacatggttaacccctgtgttccaccca 1623  
 Db 1198 CCTGCTTATACAGGTGGTGGGATGCAATATGCTGTTAAACCCCTGTGTTCACCTCA 1257  
 QY 1624 gtcttcaacaaatccaaagattgtgtcaatggttccaaatcgtcccgagtgctac 1683  
 Db 1258 GTCTTCAACCAATCCAAAGATTCTGTGTCATGCTCAAAATCGCCCCAGTGTACTAC 1317  
 QY 1684 catctgagagagtggtgtctgtatgataatgacatcgtgataacagcaaaaaaagaa 1743  
 Db 1318 CATCTGTGAGAGTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
 QY 1744 cccgtatcccttaacccctagctgtatgctcgatataagagcgagcgctgtgctgtagaaca 1803  
 Db 1378 CCGGTATTCCTTACCTTAGCTAGTATATGCTCGGATTAGGAGCGCGCTGTGGCGTATGAGACA 1437  
 QY 1804 gggacagctgcctgtatcaacagcaacagcagctagagaaggaactgtgtgactacat 1863  
 Db 1438 GGGACAGCTGCCCTGATCAGACAGGACCAAGCAGCTAGAGAAAGAGCACTTGGGAGCTACAT 1497  
 QY 1864 gggacagctgcctgtatcaacagcaacagcagctagagaaggaactgtgtgactacat 1923  
 Db 1498 GGGGCGATGACAGAGATCTCGAGCCTTAAAGAGATCTGTATACCACTTAAGAGAGTCC 1557  
 QY 1924 ctgacttcttctgtaagtggttctacagaacccgaggggattagatctgtcttctca 1983

Db 1558 CTGACTTCTTTCTGAAAGTGGTCTACAGAACCGAGGGGATGATGATGCTGTTCTTA 1617  
 QY 1984 agagaagtggtgttattgtgcacgcttaaaagaagaattgttctatagatcaacca 2043  
 Db 1618 AGAGAGGTTGGGTTATGTACAGCTTAAAGAGATGTGCTTTATATGATGATGATGAT 1677  
 QY 2044 ggaagccatagagactcactgaacagcttgaagaaagtttagagagggcgttcgaaggaa 2103  
 Db 1678 GGAGCCATGAGAGACTCCATGAAACAACTTAAAGAAAGTTAGAGAGGGGCTGAAGGGA 1737  
 QY 2104 agagaagctgacaaagggtgtgttgaaggaatgttcaacaggtctcctgtatgacacc 2163  
 Db 1738 AGAGAGGCTGACAGAGGGGCTGTTTGAAGGATGCTTCAACAGGCTCTTGGATGACAC 1797  
 QY 2164 ctgatttctgccttgacaggggccccctagatgagccctccttcttactaagttgagct 2223  
 Db 1798 CTGCTTCTGCTGTGACGGGGGCCCTAGTAGTCTGCTCTCTCTTACTTACATGTTGGCT 1857  
 QY 2224 tgcataatlaagtttctgacctgtttagagaacagatgagtgacgtccacatg 2283  
 Db 1858 TGCTTATTAATATAGTGTGTGTGCTTGTATAGAAACAGATGATGATGATGATGATGAT 1917  
 QY 2284 gtaacttagcaacagatcaacagagcctctgagccaagagagaactgactctag 2337  
 Db 1918 GTACTTAGGCAACAGATCAGGAGGCTTGTGAGCCCAAGGAGAAACGTACCTTAG 1971

RESULT 12  
 AF038601  
 Sus scrofa  
 LOCUS  
 DEFINITION  
 Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete sequence.  
 ACCESSION  
 AF038601  
 VERSION  
 AF038601.1 GI:3133304  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 pig.  
 Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 1 (bases 1 to 7333)  
 AUTHORS  
 Akiyoshi, D.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P. and Fishman, J.A.  
 TITLE  
 Identification of a full-length cDNA for an endogenous retrovirus of miniature swine  
 JOURNAL  
 J. Virol. 72 (5), 4503-4507 (1998)  
 MEDLINE  
 98216827  
 REFERENCE  
 2 (bases 1 to 7333)  
 AUTHORS  
 Fishman, J.A.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 3 (bases 1 to 7333)  
 REFERENCE  
 Fishman, J.A.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REMARK  
 Sequence update by submitter  
 On May 16, 1998 this sequence version replaced gi:3116445.  
 FEATURES  
 source  
 1..7333  
 /organism="Sus scrofa"  
 /strain="miniature swine"  
 /db\_xref="taxon:9823"  
 /cell\_line="PK-15"  
 /note="porcine endogenous retrovirus ERV-PK15; infective virus but represents a naturally occurring transcript"  
 LTR  
 598..2172  
 /codon\_start=1  
 /product="gag protein"  
 /protein\_id="AAC16766.1"  
 /db\_xref="GI:3116446"

```
/translation="MGQVTPPLSLTIDHTEVSRANHLSVQVKKPQWOTFCASENP
IDVWSESETESEIILAVAKAIIIFQTPGSHPOQEPYIITWDLADDPWPVWPN
KPRKPERIILGEMKHSKESRISRIYEIEPTEWPEPPVPPVPPVPPVPPVPPV
PSAPGAPVWVPPAGTSPRRGATPERDEIAIPLRTYGPMPGGOLOIOWWPFSS
ADLYNMKINHPFSEDPORLTGLVESLMEFSDIAPLWDCOOLLOTLPTEERETLEA
KNVYGDGRTOLONEIDMGEPFLTRGMYNTEGSEKIKYQALVAGIRGSRP
TKIAKRVPMGCPMPSPVPERLMEAFRRPTPDPSSEAKASVALAFIGOSALDIR
KKIORBIOELAEIRDVIREAEKYVYRETEEEKEREREERERRRORREKNT
KILAAVEGSSSREREDFRIRSGPROSMLGNRPDLKDOCAVCEKKGHWARNCPK
GNGKPKVLALEDK"
CDS
2320..3552
/note="truncated protein"
/codon_start=1
/product="p01 protein"
/protein_id="AAC16767.1"
/db_xref="Gen:3133305"
/translation="MGATGQROYPMTRRTVLDGVHVSFLVIPCVPYLLGRDL
TKMGALISFEGREPVSNKPKITVLTLDDERLXSPVKPRODIOSLBOPQW
AETMGALAKOVPROVIOUKASATPVROYPLSRERESIMHYQLIDOGILVPO
SPWMTPLPRKRENTNDYRYODIREVNRKVDIHPVPMYNDLSLPERMWTVL
DKDAFCFLRHPTSQLFTFEMWDPGTGRLTWTRLPQGNPTIDEALHDL
ANFRIOHPOVTLLOYVDLLAGATKODCEGRKALIELSDIGYRASAKKAOICRE
VLYLGYSLRGOSMLTEARKKTVQVQIPAPTTAKOVREPLGTAGRCRLIMPATLAAP
LYPLTKKGGCLPQGGK"
4803..6725
/note="env protein; N terminus deleted"
/codon_start=1
/pseudo
6726..7333
LTR
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
Query Match 63.7%; Score 1882.8; DB 4; Length 7333;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 402; Indels 195; Gaps 9;
QY 1 tgccttttaggttagaacaacccctggaagtttggtcgtgaacccctatgaattgctct 60
DB 4462 TGCCTTTTAGGTTAGGAGAACCCCTGACAGTTTGGCTGACCCCTATMAATTGCTCT 4521
QY 61 aacggggagacccccccgtgtgtagaattgctctgtagatgctgtagtgcgtctt 120
DB 4522 ACGGGGAGACCCCGCTGGCAGAAATGCTTGCATAGTGTGATGCTGCTCT 4561
QY 121 cccagccttctctcttagagctcaaggcgtcgtgagtgaggaacgaagcgtgagac 180
DB 4582 CCCAGCCTTGTCTCTAGCTCAAGGGGCTCGAGTGGGTGAGGACGACGAGCGTGAAGC 4641
QY 181 agcccgaggagcctactcaggaagagagaagctgcaagttccacatcgcttccagtg 240
DB 4642 AGCTCCGGAGGCTACTCAAGAG--GAGACTTGAAGTTCCACATCGCTTCCAAAGTTG 4698
QY 241 gaattcaagtctatagtagacgacgctgaggaacaccccggaacccgtgtgaaaggac 300
DB 4699 GAGATTGACGTATGTTAGACGCCACCGTGCAGAAACCTCGAGACTCGGTGAGAGGGAC 4758
QY 301 ctatctcgtacttttagaccacacaaagcgtgtgaaagtgcgaagaatctccaccgtga 360
DB 4759 CTATCTCGTACTTTTGGACACACCAACGGGCTGTGAAGTCGAAG-----4803
QY 361 tcatatcatcccaagtttaagctgagccacctccgaactcggggtgagagccgaaaaa 420
DB 4804 -----4803
QY 421 ctgagaatcccttaagctcgtccatcgtcgtgtctctactctaaacaataactccc 480
DB 4804 -----GAATCCCTTAAGCTTGGCTCCATCGCGTGTCTCTACTCTGTGAATAACTCCT 4859
QY 481 caagccagtagtaacgacctatagacagctcgaaccccatagaccttatccctacc 540
DB 4860 CAACTTAATGTAAGCCCTTGTGGAGACGCCCACTCCCATTAACCTTATCTCTCACC 4919
/translation="MGQVTPPLSLTIDHTEVSRANHLSVQVKKPQWOTFCASENP
IDVWSESETESEIILAVAKAIIIFQTPGSHPOQEPYIITWDLADDPWPVWPN
KPRKPERIILGEMKHSKESRISRIYEIEPTEWPEPPVPPVPPVPPVPPVPPV
PSAPGAPVWVPPAGTSPRRGATPERDEIAIPLRTYGPMPGGOLOIOWWPFSS
ADLYNMKINHPFSEDPORLTGLVESLMEFSDIAPLWDCOOLLOTLPTEERETLEA
KNVYGDGRTOLONEIDMGEPFLTRGMYNTEGSEKIKYQALVAGIRGSRP
TKIAKRVPMGCPMPSPVPERLMEAFRRPTPDPSSEAKASVALAFIGOSALDIR
KKIORBIOELAEIRDVIREAEKYVYRETEEEKEREREERERRRORREKNT
KILAAVEGSSSREREDFRIRSGPROSMLGNRPDLKDOCAVCEKKGHWARNCPK
GNGKPKVLALEDK"
CDS
2320..3552
/note="truncated protein"
/codon_start=1
/product="p01 protein"
/protein_id="AAC16767.1"
/db_xref="Gen:3133305"
/translation="MGATGQROYPMTRRTVLDGVHVSFLVIPCVPYLLGRDL
TKMGALISFEGREPVSNKPKITVLTLDDERLXSPVKPRODIOSLBOPQW
AETMGALAKOVPROVIOUKASATPVROYPLSRERESIMHYQLIDOGILVPO
SPWMTPLPRKRENTNDYRYODIREVNRKVDIHPVPMYNDLSLPERMWTVL
DKDAFCFLRHPTSQLFTFEMWDPGTGRLTWTRLPQGNPTIDEALHDL
ANFRIOHPOVTLLOYVDLLAGATKODCEGRKALIELSDIGYRASAKKAOICRE
VLYLGYSLRGOSMLTEARKKTVQVQIPAPTTAKOVREPLGTAGRCRLIMPATLAAP
LYPLTKKGGCLPQGGK"
4803..6725
/note="env protein; N terminus deleted"
/codon_start=1
/pseudo
6726..7333
LTR
BASE COUNT 1984 a 1788 c 1885 g 1676 t
ORIGIN
Query Match 63.7%; Score 1882.8; DB 4; Length 7333;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 402; Indels 195; Gaps 9;
QY 541 tgcgtatatttagaccctgtagcaggtgtcactgttaaatagacactcgaagtgtgctcct 600
DB 4920 TGCCTACTTACTGACCTCCGGTAGCAGTATTAAATTAATTAACGACCTCAAGGGAGGCTCC 4979
QY 601 agagagcactcgtgtggtcctgtaactcgtacttctgctccgattgattaaaccgctgtta-- 658
DB 4980 TTGGGGACCTGGTGGCTGGAATATATATGCTGCTTCGATGCACTAATTCCTGCTTCAT 5039
QY 659 -----aaagacacctcccaacccctagctcgtatgattggtctcattgtgcgaagg-----c 711
DB 5040 GACAGGCGCACACCCCGGATGATCTGCTGCTTACGGGTTTACGTTTGGCCAGACCC 5099
QY 712 acagagaagaagaatactgctgaggggtctcggggaalccctcgttagagatgagctgc 771
DB 5100 CCAATTAATGAGAAATATGTTGGAAATTCCTCAGGATTTCTTTTGAAGCAATAGGACTCG 5159
QY 772 gtcaaccctcaagatgtagagctggaatggaatggtcgaactctctccaggaacgggtaaatc 831
DB 5160 ATAACTTAAATGATGAGGAATGGAATGGCAATGGCCAGTCTCTCAGCAGACAGAGTAAGTTAC 5219
QY 832 tccctgt-----caa 842
DB 5220 TCTTTTGTAAACAATCCTACCAGTTAATCAATTTAATGCGCATGGAGAGATGAAA 5279
QY 843 ttcgagccggggaagtacaaaatgtagaaactataaaagaataagctgtccccatc 902
DB 5280 GATTGGCAACAGCGGTACAAAAGATG-TACGAATAAGCAATAAAGTGTGATTCGTT 5338
QY 903 agacttagattatcctaagaatggttctcgtgaaagaagaagaagaagaagaatcctaa 962
DB 5339 AGACTTAGATTACTTAAATAATGTTCACTGCAAGAAAAGAAAAGAAAATAATTCAAA 5398
QY 963 gtgataaatgtagtagagctgaggggaatgttttta-----ataatagcgggggagcag 1019
DB 5399 GTGGGTAATAGTATATCTTTGGGGAATAGTGTACTGTGAGAGGCTCGGAGAAAGAAAG 5458
QY 1020 gtccacttaacatcctgacctaggaatagagacgggagacacccctgtgtgcaatgg 1079
DB 5459 ATCTGTTCTGACTATTCGCTCAGATAGAAACTCAGATGAAACCTCCGGTGTGATAGG 5518
QY 1080 acccgataagtagctgctggaacaggggccccgggcccgtgggacccgataactgtcc 1139
DB 5519 ACCAAATAAGGGTTTGGCCGAACAAAGACCTCCAAATCAAGA-----5560
QY 1140 ggtgcccaatlaactcgtcgtgctgacataacacacgcccctagcaacagtagaac 1199
DB 5561 -----ACAGAGGCCATCTCCTAACCCCTGTGATTACAAATCAACCTC 5602
QY 1200 tgaattgattcctaccacaacgctagaactcccaagtgctcgttgaagacagagaca 1259
DB 5603 TGGATCAGTCCCACTGA-----GCTTACATCAGTATTAAACAGGGGC 5647
QY 1260 gagactctcagctctatcccaaggagcttccaagcacaactccaaccccgatgac 1319
DB 5648 GAAACTTTTAAAGCTCATCAGAGGACCTTTTAAAGCTCTTAACCTCAGACACTCAGAGGC 5707
QY 1320 caactctctgtgtgcttgcctatcccaaggagctccttctttagggggaatgagctaa 1379
DB 5708 TACCTCTTCTGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 5767
QY 1380 agaaaagaanaatcgaatgtagacaagaagacatagaatcaatgtagatgaggtgtccgaa 1439
DB 5768 AGGAGGGAATATCATGTGACAAAGAGACATAGAGCAATGATGAGGATGCCAANA 5827
QY 1440 taagtctaccccaatgaatgtccgggaagggagacatgacataggaanaagctcccca 1499
DB 5828 TAAGTTACCTTACTGAGAGGTTTGTGAAAGGACCTGACATAGGAGTGTCCCCATC 5887
QY 1500 ccaaccaacacttctgtagaagctgtgttagagagagcctcagaanaatcagattt 1559
DB 5888 CCAACCAACCTTTTGTAAACACATGAAAGCTTTTATGAAACCTTGTGAGAGCTATATCT 5947
QY 1560 agtactggttataacaggtgtgtggcatgacatactggttlaacccctgtgttccac 1619
```

[illegible]

Db	7028	ATCTCACTGCAGATGTGTCACCTCTCCAGAGAGCCACGACGAGATGCCGAGACT	7087
Qy	2700	attttaaatatgtatgtgtccaaagagacggggctctcgtatatttaaaatgttgtcat	2759
Db	7088	ATTTTAAATGATGTTGGTCCACGAGACGGGGCTCTCGATATTTTAAATGATGGTCCAT	7147
Qy	2760	gagacggcggcctcgtatatttaaaatgtatgtgttgacgcacaagccttcttga	2819
Db	7148	GGAGCGCGGGCTCTCGATATTTTAAATGATTTGGTTTGAACGACACAGCTTTGGTGA	7207
Qy	2820	accaccaaaaagtgtcccgatccgcacatcggggcggcagtcctctacacctgtggt	2879
Db	7208	ACCCATTAAGAGCTTCCCGATTCCGCACCTCGGGCCCGACGCTCTACCCCTGGTGGT	7267
Qy	2880	giacgactgttggcccgacggcgcgtcttggaataaaatcctctgtgtgttcataaaaa	2939
Db	7268	GTACGACGTGTGGGCCCCACGCGCCTTGGAAATTAATCTCTTGTGTTGATCAAAAA	7327
Qy	2940	aaaaa 2945	
Db	7328	AAAAA 7333	
RESULT	13		
LOCUS	ARI30474	7333 bp	DNA
DEFINITION	Sequence 2 from patent US 6190861.		
ACCESSION	ARI30474		
VERSION	ARI30474.1	GI:14118799	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 7333)		
AUTHORS	Fishman,J.A.		
TITLE	Molecular sequences of swine retroviruses method of using		
JOURNAL	Patent: US 6190861-A 2 20-FEB-2001;		
FEATURES	Location/Qualifiers		
source	1..7333		
BASE COUNT	1984 a 1788 c 1885 g 1676 t		
ORIGIN			
Query Match	63.7%;	Score 1882.8;	DB 6; Length 7333;
Best Local Similarity	80.1%;	Pred. No. 0;	
Matches 2409; Conservative	0; Mismatches 402;	Indels 195;	Gaps 9;
Qy	1	tgcttttgggttaggaacaccctctgacaagtttgggtgacccctatgaattgctct	60
Db	4462	TGCTTTTAAAGGTGAGAGAACCCCTTGACAGATTGGGCTGACCCCTATAATTGCTCT	4521
Qy	61	acggggagaccccccccgctgtgtagaatgtcttctgtacatagtgatgtgtcgtctt	120
Db	4522	ACGGGGAGACCCCCCGCTGGCAGAAATGTGCTTGGACATAGTCTGATGTGCTGCTTT	4581
Qy	121	cccaagccttgttctctatagcgtcaagggcgtctgaattgtgttgagcgaacgcttgaagc	180
Db	4582	CCCAAGCTTTGTTCTTAAGCTCAAGGCGCTGAGTGGGTGAGCAGCAGCGTGAAGC	4641
Qy	181	agctccggagagcctactactaagagagagagacttgcgaatttccacatcgcttccaagtg	240
Db	4642	AGCTCCGGAGGAGGCTACTCAGAG--GAGACTTGCAGATTTCCATATGCTCCAAAGTTG	4698
Qy	241	gagattcagttcatgttgaacgcacacccgtgcaaggaaactcgaagactcgtgtggaaggac	300
Db	4699	GAGATTCACTATGTTTAACAGCCACCCGTCAGGAAACCTCGAGACTCGGTAAGAGGAC	4758
Qy	301	cttatctgtacttlttgacacacacaacgaagcgtgtgtaaagtcgaaggaatctccacttga	360
Db	4759	CTTATCTGCTACTTTTGGACACACACACAGCGGTGTGAAGTGAAG-----	4803
Qy	361	ttcatgcatccacgltlaagcttgcgcgcacactcccgactcgggtgtgtgagagccgaaaaa	420







QY	1084	gataaagctagctgcttgaaacagggcccccgccctgagacacgcgataactctgcggtg	1143
Db	721	gataaagctagctgcttgaaacagggcccccgccctgagacacgcgataactctgcggtg	780
QY	1144	ccccaatcaactcgtcgtgcgcgcctgcacataaacacgcgccttagaagaagctacccttga	1203
Db	781	ccccaatTTAACCTCGCTGCGGCGCTGACATPACACAGCCCGCTTAGACACGGTACCACTGGA	840
QY	1204	ttgattccctaaccaacgcgccttagaanaactccccagcgtgttcctgtttaagacaggaagaaga	1263
Db	841	TTGATTCCTACCAACATGCGCTAGAAACTCCCGAGGTGTTCTGTTTAAGACAGAGAAGAGA	900
QY	1264	ctcttcagctctcatccagggagccttccaaagcatalaactccacgcgccttgaTgcacT	1323
Db	901	CTCTTCAGCTTCATCCAGGAGGAGCTTTCAAAAGCCATCAAAATCCACCGACCCGAGGCGCAT	960
QY	1324	tctctctgttggccttgcatacctcaaggccctctattagaagggaatgcttaagaa	1383
Db	961	TCCTCTGTGTGGCTTGTCTCATCTCAAGGCGCTCCCTTATTATGAGGGAGTGCTTAAGAA	1020
QY	1384	agaaatctcaatgtgaccaaagagcatalaanaatcatalaactgagagctcccgaaataag	1443
Db	1021	GGAAATTTCAATGTGACCAAGAGGAGCATAGAAATCATGTACATGGGGGTCCTCGAATATAG	1080
QY	1444	cttaccctcacTgaagtttcgggaaagggaatgatagaaagctcccccacTcccac	1503
Db	1081	CTTACCCCTCAAGGAAATTTCCGGGAGGAGACATGCATGTGAAAGAGCTCCCCATCCAC	1140
QY	1504	caaacacttgctatgtaactgtgtttatlgacagcgctccagaanaatlcagatattagta	1563
Db	1141	CACAAACCCTTGCTATATACTGTGTGTATGAAACAGGCTCGAAGAAATCAGTATTATGTA	1200
QY	1564	ccctgtcataacagctgtgtggagatbcaactactgggttaeccccctgtttccaaccca	1623
Db	1201	CCGTGTTATTAACAGGTGTGTGGCATGTCATACTGGGTTAACCCCTGTGTTTCCACCTCA	1260
QY	1624	gtcttcaaccaatccaaagaattgtgtgtcatgtgttccaatgcgtcccgagtgTactac	1683
Db	1261	GTCTTTTAAACCAATCCAAAGATTTCTGTGCATGATGTCACAAATGCTCCCGAGTGTACTAC	1320
QY	1684	catctctaggaagtggttccttgaTgaattgactatcggtataacccgcacaaagagaa	1743
Db	1321	CATCTTAGAAGAGTGCTTGATGTAATGTGATCTATCGGTATTAACCGACCGAAAGAGAA	1380
QY	1744	cccgTacccttcaaccctagctgtgaatgcTccgaattagagacgcggtttggcgtatgaaca	1803
Db	1381	CCCGTATCCCTTACCTACCTAGCTGTGAATGTCCTCGATTAGGAGCGCGCGTGGCGTAGAAC	1440
QY	1804	ggagcagctgcTccctgatcacaagagccacagcagctagagaaagaaactgtgtgaTctacat	1863
Db	1441	GGGACACCTCCCTGTATCACAGGACACAGCAGCTTAGAGAAAGGACTTGGTGTAGCTACAT	1500
QY	1864	ggggccatlgacaaagaatctccgagccttaagagcgctcgtttagaaccttagaagatcc	1923
Db	1501	GGGGCCATMGACAAAGATCTCCGAGGCTTTAGAGAGGAGTCTGTTAGACACTTGAAGAGATCC	1560
QY	1924	ctgagctcttctgtcTgaagtgtgtctacagaaaccgagagggatgtatgcTctgttctcta	1983
Db	1561	CTGAGCTCTTGTGCTGGAAGTGTTTACAGAAACCGAGGGAGTTTAGATCTGCTGTTTCTGA	1620
QY	1984	agagaagtgTgggttatgtgcagccttaaaagaagaatgttgcttctatgtatgaTactca	2043
Db	1621	AGAGAAAGTGGTATTATGTGACGCTTTAAAGAAAGATGCTTGCTTATGTAGTACATCTCA	1680
QY	2044	ggagggcatcagaaactccatlgaaacaaagcttagaaaaaagttagagagcgTcgaagggaa	2103
Db	1681	GGAGCCATYCAGAACTCCATGAGACCAAGCTTTAAGAAAGGTTTAAAGAGGCGTACGAAGGAA	1740
QY	2104	agaagcgtTaccaggggtgtgtttgaagatgtgttcaacagctctcccttgaTbaaccac	2163
Db	1741	AGAGAGGCTCACCAAGGGGTGTGTTTAAGAGATGCTTCAACAGGCTCTCTCTTGATATCACCC	1800

QY	2164	ctgctctctgctgcgaagggccccccagtagccctccctgcttacttaagctggggcct	2223
Db	1801	CTGCTTCTGCTCTGACGGGACCGGCTAGTAGTCCTGCTCTCTTACTTACGATTGGGCTT	1860
QY	2224	tgcttaataaagattgctgccttgcttagaagaagtagtgcagtcagtcagatcatg	2283
Db	1861	TGCTTAATAATATAGCTTTGTTGGCTTTGTAGGAATAGTGAATGACGATCCAGATCATG	1920
QY	2284	gtacttagggcaac	2296
Db	1921	GTACTTAGACAC	1933
RESULT	15		
LOCUS	PEN288588		
DEFINITION	Porcine endogenous retrovirus PERV-AB7 env gene for envelope protein, genomic RNA.		
ACCESSION	AJ288588		
VERSION	AJ288588.1	GI:7414445	
KEYWORDS	env gene; envelope protein.		
SOURCE	porcine endogenous retrovirus.		
ORGANISM	porcine endogenous retrovirus		
REFERENCE	Viruses; Retrovirdae; Mammalian type C retroviruses; 1-Mammalian type C virus group.		
AUTHORS	1 (bases 1 to 2000)		
TITLE	Bosch, S., Arnald, C. and Jestin, A.		
JOURNAL	Study of full-length porcine endogenous retrovirus genomes with		
MEDLINE	envelope gene polymorphism in a specific-pathogen-free large white		
AUTHORS	swine herd		
TITLE	J. Virol. 74 (18), 8575-8581 (2000)		
JOURNAL	2 (bases 1 to 2000)		
MEDLINE	Bosch, S.		
AUTHORS	Direct Submission		
TITLE	Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,		
JOURNAL	AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE		
FEATURES	Location/Qualifiers		
source	1. 2000		
	/organism="porcine endogenous retrovirus"		
	/specific_host="Sus scrofa"		
	/db_xref="taxon:61673"		
	/clone="PERV-AB7"		
	/note="French Large White SPF-pig"		
gene	1. 606		
	/gene="env"		
CDS	1. 606		
	/gene="env"		
	/note="truncated, not functional"		
	/codon_start=1		
	/product="envelope protein"		
	/protein_id="CAB86207.1"		
	/db_xref="GI:7414446"		
	/translation="MHPTLSWRHLPTRGSEPKRLRPLSFASIAMFVLTLTPQASSS		
	RLIDSSNPHRPLSLTMLIIDEDTGVTVNSRGAAPRGATWPELHFLINPAVAST		
	PMLVSYGFCCPGTEKEKYGSGSGESFCRRMSCVTISNODMWKPISLDRVKFSFV		
	SSPGYKWKMLYKMDKSCPSDLDYLTISPEKGNKRIFRSG"		
BASE COUNT	532 a 496 c 499 g 473 t		
ORIGIN			
Query Match	63.2%; Score 1868.4; DB 14; Length 2000;		
Best Local Similarity	98.6%; Pred. No. 0;		
Matches 1906; Conservative	0; Mismatches 26; Indels 2; Gaps 2		
QY	364	atgcattccacgttaagctgycgcacactcccgactcggggtgagagccgaaaagactg	423
Db	1	ATGCATCCACGCTTAAAGCTGCGCACACCTCCGAGCTCGGGGTGAGAGCCGAAAAGGCTG	60
QY	424	agaatcccttaagctgctgcctcaatgcctgattgcttactctaaataaactcccg	483
Db	61	AGAAATCCCTTAAAGCTTGCCTTCACATCGGCTGGCTTCTTACTCTTAACAATPACTCCACG	120



Mon Feb 25 07:44:31 2002

us-09-171-553b-9.feb22std.rge

---



```

|||||
401 GGAGCTGCGTCCACATCCAAACGATGAGAGCTGGAATGAGCCGATCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167
451 CAGAGCCGGGTAAATTTCTCTTGTTCATCTCCGGCCGGGACAGTACAA 500
167 SmetLysLysLysLysLysLysLysLysLysLysLysLysLysLys 184
501 AATGATGAACATATATATAAGATAGAGCTGCCATCAGACTTATGATT 550
184 yLeuLysLysSerPheThrGlyLysGlyLysGlyLysGlyLysGlyLys 200
551 AACTAAGATAGTTTCACTGAAAAAGAAAAAGAAAAAGAAAAATTTCAA 600
201 TTPLeaSerGlyMetSerTyrGlyLysValPheTyrLysTyrGlyLys 217
601 TGGATAAATAGTATGAGCTGGGGAATAGTTTATTAATATAGCGGGG 650
217 ValGlySerThrLeuThrLeuArgLysGlyLysGlyLysGlyLysGly 234
651 AGCAGGGTCCACTTACCATTCGCTTACGATAGAGAGGAGGAGAGAAC 700
234 rOProValAlaMetGlyProAspLysValLeuAlaGlyLysGlyPro 250
701 CCCCTGTGGCATGGGACCCGATTAAGTACGTGCTGAACAGGAGGCCCC 750
251 AlaLeuGlyProProHisAsnLeuProValProGlyLeuThrSerLeu 267
751 GCCCTGGAGCCACCCGATACCTTGGCGGCTGCCCAATTAACCTGCTG 800
267 gProAspLeuThrGlyProProSerAsnSerThrGlyLeuLeuProT 284
801 GCCTACATTAACACAGCCGCTAGACACATGATGATGATGATGATGAT 850
284 hAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300
851 CCACACCCCTAGAAACTCCAGGTGTTCTGTAAAGACAGGACAGAGA 900
301 LeuPheSerLeuIleGlyLysAlaPheGlyLysLysSerThrAspPr 317
901 CTCTTCACTTCATCCAGGAGCTTCCAGCCATCAACTCCACGAGCC 950
317 OAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 334
951 TGATGCCACTTCTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 1000
334 yGlyGlyMetAlaLysGlyLysGlyLysPheAsnValThrLysGlyLys 350
1001 ATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGACATAGA 1050
351 AsnGlyCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlyVal 367
1051 AATCAATGTACATGGGGGCTCCGAAATAGCTTACCTCAGTAAGTTTC 1100
367 rGlyLysGlyThrCysLysLysLysLysLysLysLysLysLysLys 384
1101 CGGAGAGGGGACATGCATAGAAAAAGTCCGCCATCCACCAACACCTT 1150
384 yTyrSerThrValValTyrGlyLysLysLysLysLysLysLysLys 400
1151 GCTATAGTACTGTGTTATGAGAGGCTCAGAAATCAGTATTAGTA 1200
401 ProGlyLysAsnArgTyrTrpAlaCysAsnThrGlyLeuThrProCys 417
1201 CTGGTTATACAGGTGGTGGGATGCAATAGCTGAGTTAACCCCTGTGT 1250
417 lSerThrSerValPheAsnGlySerLysAspLysValMetValGlnT 434
1251 TTCACCTCAGTCTTCAACCAATCCAAAGATTGTGTGATGATGTCCAA 1300
434 lValProArgValTyrTyrHisProGlyLysValValLeuAspGlyTyr 450
|||||
|||||
1301 TCGTCCCCGAGTGTACTACCATCTGAGAAAGTGTCTTGTATGATAT 1350
451 AspTyrArgTyrAsnArgProLysArgGlyProValSerLeuThrLeu 1467
1351 GACTATCGATTAACCGACCAAAAGAACCCGATATCCCTTACCTTAC 1400
467 aValMetLeuGlyLeuGlyThrAlaValGlyLysGlyLysGlyLys 484
1401 TGTATGTGTGATTTAGGAGAGGCGCTTGGCGTATGAGAACAGGAG 1450
484 lAlaLeuThrGlyProGlyLysLysLysLysLysLysLysLysLys 500
1451 CCCGATCAGAGGACACACAGCTAGAGAAAGACTTGGTGAATACAT 1500
501 AlaAlaMetThrGlyAspLeuArgAlaLeuLysGlyLysSerValSer 517
1501 GCGGCCATGACAGAAATCTCCAGCTTAAAGAGTCTTATGCAACCT 1550
517 uGlyLysSerLeuThrSerLeuSerGlyValValLeuGlnAsnArg 534
1551 AGAAGAGTCCCTGACTTCTTGTGTAAGTGTTCACAGAACCGAGAG 1600
534 lLysLeuAspLeuPheLeuArgGlyLysLysLysLysLysLysLys 550
1601 GATTAGATCTGTCTTCTTAAGAAAGTGTGTTATGTGACGCTTAA 1650
551 GluGlyCysPheTyrValAspHisSerGlyAlaIleArgAspSer 567
1651 GAAGAAATGTGCTTCTTATAGATCTCAGAGGACCATCAGAGACTCAT 1700
567 lAsnLysLeuArgLysLysLysLysLysLysLysLysLysLysLys 584
1701 GAACAAGCTTAGAAAAAGTTAGAGAGCGCTCAAGAGGAAAGAGAGCG 1750
584 spGlnGlyTrpPheGlyLysLysLysLysLysLysLysLysLys 600
1751 ACCAGGGGTGTGTTGAAGATGTTCAACAGGCTCTCTTGGATGAC 1800
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
1801 CTGCTTCTGCTCTGACGGGCGCCCTAGTAGTCTCTCTCTTACTTAC 1850
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgLys 634
1851 AGTTGGCGCTTGTCTTAAATAGGTTTGTCTCTTGTAGAGAACG 1900
634 alSerAlaValGlnIleMetValLeuArgGlnLysGlyLysLys 650
1901 TGAGTCACTCAGATCATGTACTTGAAGCAACAGTACCAAGGCTTCT 1950
651 SerGlnGlyLysThrAspLeu 657
1951 AGCCAAGGAGAAACTGACTC 1971
seq_name: gb_pat:AX002804
seq_documentation_block:
LOCUS AX002804 3482 bp DNA
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132
KEYWORDS
SOURCE
ORGANISM
porcine endogenous retrovirus.
porcine endogenous retrovirus
VIRUSES: Retrovirus; Retroviridae; Mammalian type C
retroviruses; 1-mammalian type C virus group.
REFERENCE
1 (bases 1 to 3482)
AUTHORS
Stoye,J.P. and Weiss,R.A.
TITLE
Detection of retroviral subtypes based upon envelope specific
sequences
JOURNAL
Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)
LOCATION/Qualifiers
FEATURES

```

source 1..3482  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 927 a 854 c 867 g 834 t  
 ORIGIN

alignment\_scores:  
 Quality: 3484.00 Length: 657  
 Ratio: 5.319 Gaps: 0  
 Percent Similarity: 99.696 Percent Identity: 98.478

alignment\_block:

US-09-171-553b-10 x AX002804

Align seg 1/1 to: AX002804 from: 1 to: 3482

1 MethiSProThrLeuSerTrpArgHisLeuProThrArgIlyGlyLupr 17  
 911 ATGCATCCACGTTAAGCTGGCGCCACTCCCGACTGGGGTGAGAGACC 960  
 17 OlySarGleuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34  
 961 GAAAGACATGAGATCCCTTAAGCTTCGCCATCCGCTGGTCCCTTA 1010  
 34 hLeuThrIleThrProGlnAlaSerLysArgLeuIleAspSerSer 50  
 1011 CTCTAACATAATCTCCCGAGGCACTAGTAACGCCCTTATAGACAGCTCG 1060  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 1061 AACCCCAATAGACCTTATCCCTTACCTGGCTGATTTAGACCCGATAC 1110  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgIlyThrTrp 84  
 1111 GGGTGCACTGTAAATAGCAGCTCGAGGTGCTCTAGAGGACACCTGT 1160  
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 1161 GGCCGTGAATGCAATTCCTCCGATGATTACCCCGCTGTAAAGC 1210  
 101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117  
 1211 ACACCTCCCAACCTAGTCCGTAGTATGGGTCTATGTGCGCCAGGACC 1260  
 117 rGlyLysGlyLysTrpCysGlyIlySerGlyLysSerPheCysArgArg 134  
 1261 AGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCTTCTGTAGAGAT 1310  
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 1311 GGGAGCTGCCTCACTCCAGATGGAGCTGGAATGGCCGATCTCTCTC 1360  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167  
 1361 CAGGACCGGGTAAATTCCTTGTTCATTCGCGCCGCGCAAGTACAA 1410  
 167 sMetMetLysLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184  
 1411 AGTGATGAACATATATMAAGATAGAGCTGCTCCCATCAGACTTAGATT 1460  
 184 rLeuLysLysLeuSerPheTrpGlyLysGlyLysGlnGlnLysnIleGlnLys 200  
 1461 ATCTTAAGATAGATTCTCACTGAAAAAGAAAAACAGAAAAATATTCAAAAG 1510  
 201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGlyGly 217  
 1511 TGGATTAATGGTATGAGCTGGGGAATAGTTTTATATAATATGCGGGGG 1560  
 217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThGluP 234  
 1561 AGCAGGGTCCACTTTAAACATTGCTTAGATAGAGACGGGAGACAGAAC 1610

234 rProValAlaMetGlyProAspLysValLeuAlaGlnGlnLysProPro 250  
 1611 CCCCTGTGGAGTGGGAGCCCGATAAAGTACTGCTGCAACAGGGGCCCGC 1660  
 251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267  
 1661 GCCCTGGAGCCACCGCATTAACCTTGGCGTCCCAATTAACCTGCGCG 1710  
 267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 284  
 1711 GCCGTACATACACAGCCGCTTACCAACGCTTACCTGAGGCTTCCTAT 1760  
 284 hAsnThrProArgAsnSerProGlyValProValLysTrpGlyGlnArg 300  
 1761 CCAACAGCCCTAGAAACTCCCGAGGTCTCTTATAGACAGAGACAGA 1810  
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317  
 1811 CTCTTCAGTCTCATCCAGGAGGAGCTTCCAGCATCAACATCCACCGACC 1860  
 317 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 334  
 1861 TGATGCCACTTCTTCTTGTGGCTTGTCTATCTCAGGAGCTTCCTATT 1910  
 334 yTrGlnGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisArg 350  
 1911 ATGAGGGGATGGCTAAAGAAAGAAATTCATATGACCAAAAGACATAGA 1960  
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrIleuThrGluVal 367  
 1961 AATCAATGTACATGGGGGCTCCCGAAATTAACCTTCACTCAAGATTTC 2010  
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384  
 2011 CGGAGAGGAGCATGCAATAGAGAAAGCTCCCGCATCCACACACACTTT 2060  
 384 yTrpSerThrValValTrpGlnGlnAlaSerGlnAsnGlnTrpLeuVal 400  
 2061 GCTATAGACTGTGTTATGACAGGCTCAGAAATCAGTATTTAGTA 2110  
 401 ProGlyLysAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417  
 2111 CCTGTATTAACAGAGTGTGGGCTGCAATACGAGGTTAAACCCCTGTGT 2160  
 417 lSerThrSerValPheAsnGlnSerLysAspLeuValMetValGln 434  
 2161 TTCCACTCACTCTTCAACCAATCAAAAGATTCTGTGTCATGCTCCAAA 2210  
 434 lValProArgValTrpTrpHisProGlnGluValValLeuAspGluTrp 450  
 2211 TCGTCCCGCCGAGTACTACCATCTCGAGAGGTCCTTGATGAAATAT 2260  
 451 AspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrVal 467  
 2261 GACTATCGGTATTAACGACCAAAAGAGAACCCCTATCCCTTACCTAGC 2310  
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrAla 484  
 2311 TGTAATGCTGGATTAGGAGCGGCGTTGGCTGAGAAACAGGACAGCTG 2360  
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyLeuHis 500  
 2361 CCTCATATACAGAGACACAGCAGCTAGAGAAAGACTTGGTGAAGCTCAT 2410  
 501 AlaAlaMetThrGlnAspLeuArgAlaLeuLysGlnSerValSerAsnLe 517  
 2411 GCGGCCATGACAGAGATCTCCGAGCTTAAAGAGAGTGTGTACCAACT 2460  
 517 uGlnGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534  
 2461 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTTACAGAAACGAGGAG 2510  
 534 lYleuAspLeuLeuPheLeuArgGlnGlyLysGlyLeuCysAlaAlaLeuLys 550

```

|||||
2511 GATTAGATCTGCTGTTCTTAGAGAGGCTGGTTATGTCAGACCTTAAAA 2560
551 GUGUUCUCySpHeTyValaAspHisSerGlyAlaIleArgAspSerme 567
2561 GAAGAATGTTGCTTATGATGATCACTCAGAGGACCATAGACACTCCAT 2610
567 taaNlySLeuArGlyLysLeuGluArGArGArGArGArGluAlaA 584
2611 GAGCAAGCTTAGAGAAAGGTTACAGAGGCTCGAAGGAAAGAGAGCTG 2660
584 spGInGlyTTPheGInGlyTTPheAsnArGSerProtrHeThmTr 600
2661 ACCAGGGGTGTTGAAGGATGTTCAAGAGGCTCTCTGGATGACCACC 2710
601 LeuLeuSerAlaLeuThrGlyProLeuValIleLeuLeuLeuLeu 617
2711 CTGCTTTGCTGCTGACGGGACCCCTAGTAGCTCTGCTCTGTTACTTAC 2760
617 rValGlyProCysLeuIleAsnArGpHeValAlaPheValArGluArG 634
2761 AGTTGGGCTTGGCTTAATTAAGTTGCTGCTTTGTAGAGAACGAG 2810
634 alSerAlaValGlnIleMetValIleArGInGlnIleGlnIleu 650
2811 TGAGTGCATCCAGATCATGTAATTAGGCAACAGTACCAAGGCTCTTG 2860
651 SerGInGlyGluThrAspLeu 657
2861 ACCCAAGAGAAAGTACGCTC 2881

seq_name: gb_v1:PERENV2

seq_documentation_block:
LOCUS PERENV2 3482 bp RNA VRL 14-OCT-1997
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.
ACCESSION Y12239
VERSION Y12239.1 GI:2576324
KEYWORDS env gene.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
Virus: Retroviridae; Retrovirinae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 3482)
Le Tissier, P., Stoye, J.P., Takeuchi, Y., Patience, C. and Weiss, R.A.
Two sets of human-tropic pig retroviruses
NATURE 389 (6652): 681-682 (1997)
JOURNAL 97478526
REFERENCE 2 (bases 1 to 3482)
Stoye, J.P.
Direct Submission
Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
REMARK Location/Qualifiers
FEATURES
source
1..3482
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/label="PERV-B"
911..2884
/gene="env"
911..2884
/gene="env"
/codon_start=1
/protein_id="CAA72928.1"
/db_xref="GI:2576325"
/translation="MHPTLSMRHLPTRGSEPKRLRIPISFASIAMFLITLTPQASK
RLIDSSNPHRPLSLTWLIDPDTGVNSTRVAGRWMPPLHLCRLINPASTP
PNLIVSYGCGPGTEKEKYGCGSGESGCRMSCVTSNDGDMKKMPSLQDVKKPSFN
SGPKYKVMKLIKDKSCSPSIDYLIKISPTREKRENTOKNGSKGIVTKGGCA
GSTLIRLRIETGTTPPAVGVKDLAEGGPALPEPPLPPLTSLRPDTPPSN

```

```

GTTGLIPTNPRNSPGVPKGTGRLFSLIGAFQALNSTDPATSSCMLCLSSGPVY
EGMAKKGKFNVRKEHNNOCCTMGSRNKLTLTEVSGGTCTIGKAPSHOHLCSTVYEO
ASNOYLGVGNRMWACNTGILPCVSTSVFNSKPCVAVQVPRYTHPREVVIDET
DYRNNPKRKPSPSLTIAVMGLGTAVGVTGTAALTGPOULEKGLDHAAMTDLR
ALBESVSNLEESITSLSEVLQNRKGLDLTFREGLCAALKECCFVYDHSGLRDS
MSKLRERLRERREARDGMEGWRNRSNPMWTTLSALTGLVLLVLLLTLPCLIN
RFVAFVREVRVSAVOIVLHKOQYQGLISQGETDL"
BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN

alignment_scores:
Quality: 3484.00 Length: 657
Ratio: 5.319 Gaps: 0
Percent Similarity: 99.696 Percent Identity: 98.478

alignment_block:
US-09-171-553b-10 x PERENV2 ..
Align seg 1/1 to: PERENV2 from: 1 to: 3482

1 MethisProthrLeuSerTrpArgHisLeuProthrArgGlyGlyIupr 17
|||||
911 ATGATCCACAGTTAACTGAGCTGGCCACCTCCGACTGGGATGAGAGCC 960
17 OlyArgLeuArGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
961 GAAAGACTGAGATCCCTTAACCTCGCTCATGCTGCTGCTGTTCTTA 1010
34 hrlLeuThrIleThrProGlnAlaSerSerIlyArGlyLeuIleAspSer 50
|||||
1011 CTCTAACAAATAACCTCCAGCCAGCTAGTAAGCCCTTATAGACACTCG 1060
51 AsnProHisArGProLeuSerProThrTrpLeuIleIleAspProAsp 67
|||||
1061 AACCCCATACACCTTATACCTTACCTGCTGATATATGACCTGATAC 1110
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
1111 GGGGTGACAGTAAATAGCACTGAGAGTGTGCTCCAGAGGACACTGCT 1160
84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
1161 GGCTGACACGATTTCTGCTCCGATGATTAACCCCGCTGTAAAGC 1210
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
|||||
1211 ACACTCCCAACCTAGTCCGATGATGATGATGATGATGATGATGATG 1260
117 rGluLysGluLysTrpCysGlyLysSerGlyLysSerPheCysArgArg 134
|||||
1261 AGAGAAAGAAATACCTGCTGGGCTTCTGGGATCTCTCTAGAGAGAT 1310
134 rPercyValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
|||||
1311 GGAAGTGGCTGACCTCCCAAGATGAGACTGGAATAGCCGCAATCTCTC 1360
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
|||||
1361 CAGGACCGGGTAATTTCTCTTGTCAATTCGGCGCCGGCAAGACAA 1410
167 sMetIlyLeuTrpLysAspLysSerCysSerProSerAspLeuAsp 184
|||||
1411 AGTGATGAACATATATAAGATAGAGCTGCTCCCATAGACTTAGATT 1460
184 rLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 200
|||||
1461 ATCTAAAGATAAGTTCACTGAAAAAGCAAAACAGAAATAATTTCAA 1510
201 TrpIleAsnGlyMetSerTrpGlyIleValIlePheTrpLysTrpGly 217
|||||
1511 TGATATAAGGTATAGAGCTGGGGAATAGTTTATTAATATAGCGGGG 1560

```



217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 234  
 1561 AGCAGGGGTCACCTTAACCATTCGCCCTTAGAGAGACGGGACACAAAC 1610  
 234 rOProValAlaMetGlyProAspLysValLeuAlaGluGluGlyProPro 250  
 1611 CCCCTGTGGCAGTGGGAGCCGATTAAGACTGCGCTGAACAGGGGCCCCCG 1660  
 251 AAlaLeuGluProProHisAsnLeuProValProGluLeuThrSerLeuAr 267  
 1661 GCCCTGGAGCCACCGCATTAACCTCGCGTGCCTCCCAATTAACCTCGCGC 1710  
 267 gProAspLeuThrGluProProSerAsnSerThrThrGlyLeuLeuProT 284  
 1711 GCGTGCATTAACACAGCCGCTACCAAGGATCACCATGATGATTCCTTA 1760  
 284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGluArg 300  
 1761 CCACACGCCCTAGAAACTCCCGAGTGTCTCTGTTAAGACGACGACAGAGA 1810  
 301 LeuPheSerLeuIleGluGlyAlaPheGluAlaIleAsnSerThrAspPr 317  
 1811 CTCTTCATCTTCATCCAGGAGCTTTCACACCATCAATCCACCGACGCC 1860  
 317 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 334  
 1861 TGATGCCACTTCTCTTCTTGTGGCTTGTCTATCTCTCAAGGCGCTCTATT 1910  
 334 yrgGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350  
 1911 ATGAGGGGATGGCTAAAGAGAAATTCATGTGACCAAGACATAGA 1960  
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367  
 1961 AATCAATGTACATGGGGGTGCCGAATTAAGCTTACCTCCATGAGTTTC 2010  
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384  
 2011 CGGAGAGGGGACATGATAGGAAAGCTCCCGCATCCACACACACCTTT 2060  
 384 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400  
 2061 GGTATAGTACTGTGTATGAGCAGCGCTCAGAAATACATATTTAGTA 2110  
 401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417  
 2111 CCGTGTATTAACAGGTGGTGGCATGCAATACTGGGTTAAACCCCTGTGT 2160  
 417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434  
 2161 TTCCACCTCAGTCTTCAACCAATCCAAAGATTTCGTGTACATGGTCCAA 2210  
 434 lValAlaProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 450  
 2211 TCGTCCCCCGAGTGTACTACATCTCTGAGGAAGTGTCTGTGATGAATAT 2260  
 451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467  
 2261 GACTATCGGTATTAACCGACCAAAAAAGAACCCGATCCCTTACCTACGC 2310  
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484  
 2311 TGTAAATGCTCGGATTAGGAGCGCGTGGCTAGGACAGGACAGACAGCTG 2360  
 484 lAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 500  
 2361 CCTGTATCAACAGCACACGACGCTAGAGAAAGACTTGGTAGCTACAT 2410  
 501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517  
 2411 GGGGCCATACAGAAAGATCTCGACGCTTAGAGAGCTGTGTAGCAACCT 2460  
 517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534

2461 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTTCTACAGAACCGAGG 2510  
 534 lYLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaIleLeuLys 550  
 2511 GATTAGATCTGCTGTCTTAAGAGAGGTGGGTATGTGCACGCTTAAAA 2560  
 551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerme 567  
 2561 GAAGAAATGTTGCTTCTATGTAGATCTCAGAGACCATAGAGACCTCAT 2610  
 567 lAsnLysLeuArgLysLysLeuGluArgArgArgArgGluArgGluAla 584  
 2611 GACCAAGCTTAGAGAAAGGTTAGAGAGCGCTCGAAGGAAAGAGAGCTG 2660  
 584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrTr 600  
 2661 ACCAGGGGTGGTTGAAGATGGTTCACACAGGTCTCTGGATGACCCACC 2710  
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 2711 CTCTCTTCTCTCTGACGGGACCCCTAGTAGTCTCTCTCTGTACTTAC 2760  
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 634  
 2761 AGTTGGGCTTGGCTTAATTAATAGTTGTTCCTTGTAGAGAACGAG 2810  
 634 aLserAlaValGlnIleMetValLeuArgGlnGlnIleTyrGlnGlyLeu 650  
 2811 TGAATGAGTCCAGATCATGATGCTTAAGGCACAGTACCAAGGCTTCTG 2860  
 651 SerGlnGlyGluThrAspLeu 657  
 2861 AGCCAAGAGAAACTGACCTC 2881

seq\_name: gb\_v1: PEN293657

seq\_documentation\_block: 8763 bp DNA VRL 30-MAY-2001  
 LOCUS PEN293657 8763 bp DNA VRL 30-MAY-2001  
 DEFINITION Porcine endogenous type C retrovirus proviral gag gene, pol gene  
 and env gene, class B, clone 213.  
 ACCESSION AJ293657 GI:14275841  
 VERSION AJ293657.1 GI:14275841  
 KEYWORDS env gene; envelope; gag gene; group specific antigen; pol gene;  
 polymerase; polyprotein.  
 SOURCE porcine endogenous type C retrovirus.  
 ORGANISM Viruses; Retroid viruses; Retroviridae; Gamma retrovirus.  
 REFERENCE 1 (bases 1 to 8763)  
 Krach, U., Fischer, N., Gzauder, F., and Tonjes, R.R.  
 Comparison of replication-competent molecular clones of porcine  
 endogenous retrovirus class a and class b derived from pig and  
 human cells  
 J. Virol. 75 (12), 5465-5472 (2001)  
 JOURNAL 21256017  
 MEDLINE 2 (bases 1 to 8763)  
 REFERENCE Toenjes, R.R.  
 Direct Submission  
 TITLE Submitted (31-Jul-2000) Toenjes R.R., Medical Biotechnology,  
 JOURNAL Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-95, Langen, 63225,  
 GERMANY

FEATURES  
 source location/Qualifiers  
 1..8763  
 /organism="porcine endogenous type C retrovirus"  
 /proviral  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:70540"  
 /clone="213"  
 /cell\_line="PK15"  
 /note="class B"  
 1..630  
 467  
 /note="Cap site"

```

primer_bind 633..650
gene 1077..2651
/ gene="gag"
CDS 1077..2651
/ gene="gag"
/ product="group specific antigen"
/ protein_id="CAC39625.1"
/ db_xref="GI:14275842"
/ translation="MGQVTPPLSLTLDHWTENKRAHNLVSVYKKKPMQTFCASEMP
TFDVGWSEGFENSEIILAVKALIFQTPGSHPOEYILTMQDLAEPPWPMLN
KPRKCPRIALAEKKNHSAKEVSPGIYEIEEPTEPPEPPVPAQAVRG
PSAPGAPVVEGPAAGTSSRGATPERTDELILPLTGPMPAGLOLOQWPPSS
ADLYNMKTNNPPESEDPORLGLVESLMEHPQWDCOOLDTLTTERERILLEA
RKVYRGADGRPTOLQNETIDMGFLTRPGWDYNTAEGRSLKIRQALVAGLKSARP
TNLAKREVWQGPNEPPSVLEKLEMAFRFPDPPTSEAKASVALAFIGOSALIR
KTLQRLGLOEALRDLREAEKYYRRETEKEQKEREERERDRRQENKIL
KIIAAVEGKSSRERDRFRKIRSGPQSGNLNRPFLDKQCAVCYCKEKHMANRCPK
KGNKGPVLALEEDKD"
2652..6239
/ gene="pol"
/ gene="pol"
/ note="complete pol ORF, gag-pol is expressed as a
polyprotein"
/ citation=[1]
/ codon_start=1
/ product="polymerase"
/ protein_id="CAC39626.1"
/ db_xref="GI:14275843"
/ translation="GRRGSDPLPEPRVTLKVEGQVPELVDTGAHSYLLQPLGRIKE
KKSWMGATGQROYPTTRRTVDLAVGRVTHSLVIECPVLLGHDLLTKGAQISF
EGRPVSVNNKPIVTLQDLDEYRLYSQVYPPDDIQSMLQEPQVQAEATGAGMLA
KQVPPVQTLKASATPVSVQRYLSEAREGQWIPVQRLIOGILVPPVSPWNTPLP
VRPGNDYRVPQDLREVNKRVODIHPVNPENLISALPEPENNVTVDLDLAEFCI
RLHPTSQPLFAEFMRDQPGTGQTLFTWLPPOGKNSLPPEPEALHROLDRIQHPQ
VTLLQYVDLLAGATKODCLBGTALLLEISQIGRAKAKKQICRRVYTLGSLK
GGORMLTEARKKTVVOIPATTKAKOYREFLTAGFCRLMIPGATVLAFLYLTKEGK
ERSMABEHOKAPDAIKALISAPALAPVTKPEFLYDERGVANAGVLTQILGMRK
PVAYLSKRLDPVAVSGMPVCLKAIAAVALIKADAKLTLCQNTIVLAHLENIQVOP
DRMWTNARMTVOSLLTERVTEFAPALNPATLLPEETDEVTGHCQHLLEETGVR
KDTLDPLTGEVLTWPTDGSYVEGKMAAGVAVDQRTIASSIPRESTGAKRMEI
ALTOALRLAEKSIINYDSTRYAFATAHVGAIVYQGGILTSAGRIKKEISILE
ATHLPRKALIHCPGHOKAKDLISGNQADYVAKQAQAVNVLPIETPKAPEPRQ
YTIEMQELIKIDQSEETEGCTIYSKEILIPKEGLEYVOQIHLHLGTHLQOL
VRSPTVHLRLPQVADVSVKHCVPQOLVANPSRIIPGKRLGSHGAMVEDFTEVK
PAKYGNKYLLFVVDYFESGWEAYPTKETSIVAKKILEEIPPRGIPKVISGNGPA
FAVQVSOGLAKILGIDMKLCAYRPOSSGOVRMNTIKETLTKLTETGIDMTAIL
PFLVLEFRNTPGOGLTPVELLYGGBPLVETASVHSDVLLSOPLESLKALEVRQ
RAMKOLREAYSGEDQVPHRYQVDSYVRRHHHGLETFRKKGFLYLLTTPYAVKY
EGISTWTHASHVKLAPDPDSGWRAEKTEMLRLRLRPVSYNNNSPQO"
6112..8085
gene 6112..8085
/ gene="env"
/ gene="env"
/ gene="env"
/ codon_start=1
/ product="envelope"
/ protein_id="CAC39627.1"
/ db_xref="GI:14275844"
/ translation="MHPILSWRHLPTRGGEPRRLRIPISASIAWFLTLITTPAASK
RLIDSNPRLSLIWLIIIDPTGVTNSTRVAVRGMWPELHCLILINPAKSTP
PMLVSYGVCOPPEKTEKCCGSGESGSRSCSTSDGDMKMTISLQDRKFSFN
SGPGKTKWKKLKDKSCSPSDLDYLIKISFTEKGKENTQKINSGMSIVFYKYGGA
GSTLIRLKEIETGEPVAVGPDVKLAEOGPALPEPNLIPVPLTSLRPDITPYSX
STTGILPNTPRNSPGVAVKQGLRSLIOGAFOALINSTDDATSCMLCLSSGPPY
EGMAKEGFENVTKERNOCSTRNKLFLTEFSGGCTIGAPSHOHLCINSTVYED
ASENOYLVGYNRMWACNGLPCVSTSVFNOCKFOCMQOIVPVVYVHPREYVDEY
DYRYRPPKREPESLTLAVMLIGLGTAVGCTAALITIPQLEKGLGHLHAAEMDEL
ALEEVSUNLESITSLSEVYLNQNRGLDLRLRBSGLCAALKERCCFVHDSGAIRIS
MSKULERLERRERREADQGWEEGWNPSWMTLLSLATGPIVLVLLLTLLTVGCLLIN
RFAVAFREVRVSAVOIMLRQOYOGILLSQETDL"
8133..8763
LTR 8133..8763
polyA_signal 8655..8660

```

```

BASE COUNT 2301 a 2200 c 2287 g 1974 t 1 others
ORIGIN
alignment_scores:
  Quality: 3472.00
  Ratio: 5.317
  Percent Similarity: 99.391
  Percent Identity: 98.326
alignment_block:
US-09-171-553B-10 x PEN293657
Align seg 1/1 to: PEN293657 from: 1 to: 8763
1 MethisProthruserTriPARHISleuProthrusarglyglyluPr 17
|||||
6112 ATGATCCACAGTTAAAGCTGGCGCACCCTCCGACTGGGCTGAGAGGCC 6161
17 OlyargleuArgIleProleuSerPheAlaSerIleAlaTriPheLeu 34
|||||
6162 GAAAGACTGAGATCCCTTAAGCTTCGCTCATCGCTGCTCCTTA 6211
34 hrieuthrIlethrProGlnAlaserIysargleuIleasPserSer 50
|||||
6212 CTCTAACAAATAACCTCCAGGAGCTAGTAACGCCCTTATAGACACTCG 6261
51 AsnProhisArgProleuSerProthTriPheIleleasProhisprh 67
|||||
6262 AACCCCATGACCTTTATCCCTTACCTGCTGCTATTTATGACCTGATAC 6311
67 rglValThrValAsnSerThrarglyValAlaProarglyThrTrpT 84
|||||
6312 GGGTGCTACGTAAATAGCAGCTGAGGTGTGCTCCAGAGGACCTGGT 6361
84 rPrroGluLeuHisPheCysLeuArgleuIleasnProAlaValIysSer 100
|||||
6362 GGGCTGACACGATTTCTGCTCGATGATTAACCCGCTGTATAAACG 6411
101 ThrProGlnleuValArgSerTrglyPheThrCysCysProGlyTh 117
|||||
6412 ACACCTCCACACTAGTCCGTATTGCTTATTGCTGCCAGGACGAC 6461
117 rGluLySGlyLysTrCysGlyIysSerGlyIysSerPheCysArgArgT 134
|||||
6462 AGAGAAAGAAATACTGTTGGGGGCTCGGGAATCTCTGAGAGAGAT 6511
134 rPserCysValThrSerAsnAspGlyAspTrIysTrProIleSerIe 150
|||||
6512 GAGCTGCGGCACCTCCACGATGAGACTGGAAATGGCCGATCTCTCTC 6561
151 GlnAspArgValIysPheSerPheValAsnSerGlyProGlyLysTrly 167
|||||
6562 CAGACCGGGTAAATTTCTCTTTGTCATTTCCGGCCCGGCAAGACAA 6611
167 smetMetLysLeuTyrlLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
6612 AGTGATGAATAATATATAAGATGAAGCTGCTCCCATCAGACTAGATT 6661
184 yIleuLysIleSerPheThrGlyLysGlyGlnGlnAsnIleGlnLys 200
|||||
6662 ATCTAAAGATTAAGTTTCACTGAAAAAGAAAACAGAAAATATCTCAAAAG 6711
201 TrpIleAsnGlyMetSerTrpGlyIleValIlePheTyrlLysTrglyGly 217
|||||
6712 TGGAATAAGTATGATGCTGGGGAATAGTTTTTTTAAATATATGCGGGGG 6761
217 yAlaGlySerThrLeuThrIleargLeuArgIleGluThrGlyThrGluP 234
|||||
6762 AACAGGCTCACTTAACCATTCGCTTAGATAGAGACGGGAGACAGAAC 6811
234 rObProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
|||||
6812 CCCCTGTGCACTGGGACCCGATTAAGTACGTGAGAACAGGGGCCCGC 6861

```

251 AlaLeuGluProHisAsnLeuProValProGlnLeuThrSerLeuArg 267  
 |||||  
 6862 GCCCTGAGCCACCGCATTAACCTGGCGGTGCCCAATTAACTCCCTGGC 6911  
 267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 284  
 |||||  
 6912 GCCCTGACATACACAGCCGCTTAGCANCAGTACACTGGATGATTCCTA 6961  
 284 hrAsnThrProArgAsnSerProGlyValProValIleThrGlyGlnArg 300  
 |||||  
 6962 CCACACGCCCTAGAAACTCCCGAGTGTCTCGTTAGACAGACAGACAGA 7011  
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317  
 |||||  
 7012 CTCTCATGCTCATCCAGGAGCTTTCACAGCATCATCTCCACCGACCC 7061  
 317 oAspAlaThrSerSerCysThrLeuCysLeuSerSerGlyProProTyrT 334  
 |||||  
 7062 TGATGCCACTTCTTCTTGTGGCTTGTCTATCCTCAGGGCCCTTATT 7111  
 334 yrlGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350  
 |||||  
 7112 ATGAGGGGATGCTAAAGAAAGAAATCAATGATGACCAAGAGCATAGA 7161  
 351 AsnGlnCysThrThrProGlySerArgAsnLysLeuThrLeuThrGluVal 367  
 |||||  
 7162 AATCAATGATACATGGGGGTCCGAAATTAAGCTTACCTCATGAACTTTC 7211  
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384  
 |||||  
 7212 CGGGAAGGGACATGCATAGAAAGCTCCCGCATCCACACACACTTT 7261  
 384 ysfYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400  
 |||||  
 7262 GCATAGTACTGTGTTATAGACAGAGCCCTCAGAAATCAGTATTAGTA 7311  
 401 ProGlyLysAsnArgTrrProAlaCysAsnThrGlyLeuThrProCysVal 417  
 |||||  
 7312 CCGGTATTAACAGGTGGTGGCATGCAATACGTGGGTAAACCCCTGTGTG 7361  
 417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434  
 |||||  
 7362 TTCACCTCATCTTCAACCAATCCAAAGATTTCTGTGTCATGCTCCAA 7411  
 434 lValProArgValTyrTyrHisProGluGluValIleAsnArgLysTyr 450  
 |||||  
 7412 TCGTCCCCCGAGTGTACTACCATCCGAGAGGTGCTTGAATAT 7461  
 451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467  
 |||||  
 7462 GACATATGCTATACCCACCAAAAAGAAAGAACCCCTATCCCTTACCTAGC 7511  
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 484  
 |||||  
 7512 TGTATGCTCGGATTAGGAGCGGCGCTTGGCTAGGAACAGGAGACAGCTG 7561  
 484 lAlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyLysLeuHis 500  
 |||||  
 7562 CCCGATCACAGGACAGCAGCTAGAGAAAGACTTGTGAGCTATCAT 7611  
 501 AlaIleMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 517  
 |||||  
 7612 GCGGCGCATGACAGAGATCTCCGAGCTTAGAGAGAGCTGTGAGCAACCT 7661  
 517 uGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534  
 |||||  
 7662 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTCTACAGAACCGGAGGG 7711  
 534 lLysLeuAspLeuLeuPheLeuArgGluGlyLysGlyLeuCysAlaIleLeuLys 550  
 |||||  
 7712 GATTAGATCTGCTGTCTTAAGAGAGGTGGTATGTGACAGCCTTAA 7761

551 GluGluCysCysPheTyrValAspHisSerGlyValAlaIleArgAspSer 567  
 |||||  
 7762 GAGAGATTTCTTCTATGTATGATCCTCAGAGACCACTCAGAGACTCAT 7811  
 567 lAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaA 584  
 |||||  
 7812 GAGCAGACTTAGAGAAAGTTAGAGAGCGCTCGAAGGAGAAAGAGAGCTG 7861  
 584 spGlnGlyTrrPheGluGlyTrrPheAsnArgSerProTrrMetThr 600  
 |||||  
 7862 ACCAGGGGTGCTTGAAGAGATGTTCAACAGCTCTCTTGGATATACACC 7911  
 601 lLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 |||||  
 7912 CTGCTTCTGCTGACGGGGCCCTAGTACTCCGCTCGCTTACTTAC 7961  
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 634  
 |||||  
 7962 AGTTGGGCTTGTAAATTAAGTTGTGCTTGTGTAGAGAACGAG 8011  
 634 aLserAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650  
 |||||  
 8012 TGAGTGCAGTCCAGATCATGCTACTTAGGCAACACTACCAAGGCTTCTG 8061  
 651 SerGlnGlyGluThrAspLeu 657  
 |||||  
 8062 AGCCAAGAGAAACCTGACCTC 8082  
 seq\_name: gb\_v1:PERY17013  
 seq\_documentation\_block:  
 LOCUS PERY17013 7808 bp DNA VRL 14-APR-2000  
 DEFINITION Porcine endogenous retrovirus gag, pol, env genes.  
 ACCESSION Y17013.1  
 VERSION Y17013.1 GI:3005985  
 KEYWORDS env gene; gag gene; pol gene.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM  
 Viruses: Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 7808)  
 CAuderns, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.  
 Establishment and characterization of molecular clones of porcine  
 endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)  
 MEDLINE 20219394  
 JOURNAL  
 TITLE  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 SOURCE  
 1..7808  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 /tissue\_type="kidney"  
 /cell\_line="PK15"  
 /gene="gag"  
 /codon\_start=1  
 /protein\_id="CAA76581.1"  
 /db\_xref="GI:3005985"  
 /db\_xref="SPTREMBL:O73504"  
 /translation="MGQVTVTPLSITDIDHMTVEVSRRAHNLVYKKKPMOTFCASEMP  
 TEDGWPSEGFNSFETLLAKAITFOGRPSHPOPEYILTWDLAEDPPAPYKPMIN  
 KPRAPGRILLALGKNKHSKEVPSRITPELEPTPEPPVPPPIPAOGAVRG  
 PSAPGAPVVEGPAAGTASRGATPEPTDEIALLPLTGPMPGGQLDPLQTPSS  
 ADLVNMTNHPPESEDPORLGLVSEIMFESHOPWDCOOLQTLFTAERERILLEA  
 RKNVAGADGPTOLNEDMGFPLTRIGMDYNTAEGRSLKITYRQALVAVLGRASRP  
 TNLAKREVWGQPNPEPSVPLERIMEAFRRPTPDPTSEAKSVALAFVIGOSALDIR  
 KTLQRLGLODAELRDILVREAEKAYIGRETBEEKQKKEREERERDRRQREKNT

gene  
CDS

gene  
CDS

```

KILAAVVEGKSSRRERDFKRIKINSGPSQGNLGNRPDLKQDCAVCEKKGHMANRCPK
KGNKGPVLALEEDKD*
2120..4726
/gene="pol"
2120..4726
/gene="pol"
/codon_start=1
/protein_id="CAAF6582.1"
/db_xref="GI:3005987"
/db_xref="SPTREMBL:073505"
/translation="MGAOISFEQGRPEVSNNKPITVLTLOLDDEXEATLSPQKPPQD
IOSLLEQFQAMAEAGMGLAKPPVPOVTOLKASATPVSRQYPLSREAREGIPVQ
RLIOGGLIVYOSFMTPLPRKPGINDRPVQDLREVKRVQDILHPVNPYNLLS
ALPERNNWTVLDLDAFCLRLHPTSPPLAEEMRDPGRTGOLWTLRDPGFNS
PTLEDEALHRLDANFRLOHPQVTLQYVDLLLAGATKOPCLGTLLELSDLYR
ASAKKQICREYVLYGSLRGQRMLEAKKTVQVQIPAPYAKOVBREFTAGFCR
LMIKGFATLAAPLYPLTEKEGSESNAPBEHOKADAIKKALLSPALADPVPKPLY
VDKRGYARGLVLOTLOQPMRPRVAYLSKRDIPASGMPCLKIAAVALLVKQDKLT
LGNITVYAHALENYKOPPRKMTNARHTHOSLTERVFAFPALINPATLLE
ETDEPVTHDCHOLLIEFTGVKDLTDIPLTEVLWETDGSYVVGKRNAGAAVVG
THTIMASSLEPGTSAQKAEMLALTOALRLAEGKSINIYDSRYAFATVHGAIVYKQ
GLTSAGREIKNEEILSLLEALHPKRLAIHCPGHOKAKDLISNGMADRVAKQA
AOAVNLPLIETPRAPEROVLTEDMOETKKIDQSETEGTCYSXKEILPHKEG
LEYVOOIHRLTHLCTKHIOQVLTSPYHVRLEGVADVYKHCVPQOLYNANPSRILP
GKRLRGSHBGAHWEVDPTKYPKAKYGNKILVFDVTSQWKEAYPRKKESTVAKKY
WRKFFRDLERL*
5294..7267
/gene="env"
5294..7267
/gene="env"
/codon_start=1
/protein_id="CAAF6583.1"
/db_xref="GI:3005988"
/db_xref="SPTREMBL:073506"
/translation="MHTLSMRHLPTRGGEPRKRLRPLSFASIAWFLTLTTPQASX
RLIDSSNPHRPLSLTWLTIIDPRLGVNSTRGVAPRGTMWPELHCLRLINAVKSTP
PILVRSYGFCCGTGTEKREKYGSGSESFCRRNSCVTSNDGDKWPISLDDRYKFSVN
SGRGKRYKWKLYKDKSCPSDDLYLKTISTEFGKQDNIQKWMNGMSWGLVFYKYGGA
GSTLTLRLKLETGTEPPVAVGDPKVLAEQGPALPEPHNLVPOLTLSPDITQPSN
GTTGLIPTPTNPSGVPVKTQQLRFLSLQGAFOAINSTDPAISCMWLSGPPY
EGMAKGEKNVTKREHNOCTGMSRNKLTLEYSKGTCIGKAPSPHOLHCYSTVYEO
ASENOYLIVGYNNMAMCNGTLPCVSSSVNOKPCVWQVQIIPVYXYREPVVDEY
DYRNRKREPVSITLAVMLGCTAGVGTCAALITGQQLKEKIGELHAAMTDLR
ALEESYNSLESISTSEVYLVONRRGLDLFLREBGLCALKECCFVYDHSGLRDS
MSKRLERLRRRRERADGWEFGWPNRSMWTTLSALTGLVLLTLITVGPCLIN
REVAFEREVGVAVQIMVLNROQYQGLSGSETDL"
BASE COUNT      2077 a      1971 c      2047 g      1713 t
ORIGIN

```

alignment\_scores:  
Quality: 3468.00 Length: 657  
Ratio: 5.303 Gaps: 0  
Percent Similarity: 99.543 Percent Identity: 98.021

alignment\_block:  
US-09-171-553B-10 x PERY17013 ..

Align seg 1/1 to: PERY17013 from: 1 to: 7808

```

1 MethisProthleuSerTrpArgHisleuProthtrArglygluIupr 17
|||||
5294 ATGCATCCACGTTAAAGCTTGGCGCCACCTCCGAGCTCGGGGTGAGAGGCC 5443
17 OlyArgleuArgileProleuSerPheAlaSerIleAlaTrpPheleut 34
|||||
5344 GAAAGACTGAGANTCCCTTAAGCTCGGCTCATGCTGCTGCTCTTA 5393
34 hileutIlelethProglAlaSerSerlyArgleuIleAspSerSer 50
|||||
5394 CTCTAACAAATAACCTCCAGGACAGTAGTAAGCCCTTATAGACAGCTCG 5443
51 AspProthIsArgProleuSerProthtrPheleuIleleAspProksph 67

```

```

|||||
5444 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTTATGACCTGTATAC 5493
67 rglYalThrValAsnSerThrArglyValAlaProArglyThrTrp 84
|||||
5494 GGGTGTCACTGTAAATAGCACTGAGGTGTTCCTTAAGAGGACCTGCT 5543
84 rPProglIleuHisPheCysleuArgleuIleAsnProAlaValIlySer 100
|||||
5544 GGCCTGAACATCATTTCTGCTCCGATTGATTAACCCCGCTGTTAAAGC 5593
101 ThrProAsnleuValArgSerTrpGlyPheTrpCysCysProglIly 117
|||||
5594 ACACCTCCCAACCTAGTCCGAGTATGGGTTCATATTCGCCAGGAC 5643
117 rglIlyGlyIlySerTrpCysGlySerGlyIlySerPheCysArgArg 134
|||||
5644 AGAGAAAGAGAAATAGTGTGGGGTCTGAGGAATCCTCTGTAGAGAT 5693
134 rPserCysValThrSerAsnAspIlyAspTrpIlySerProIleSerleu 150
|||||
5694 GGACCTGCTGCTACCTCCACGATGAGACTGGAATGGCCGATCTCTC 5743
151 GluAspArgValIlyPheSerPheValAsnSerGlyProglIlyTrp 167
|||||
5744 CAGGACCGGCTAAATATCTCTTGTCAATTCGGCCCGGCAAGTAA 5793
167 smetMelyleuIlyTrpIlyAspIlySerCysSerProSerAspLeuAsp 184
|||||
5794 AGTAGTAACACTATATTAAGATAGAGCTGCTCCATCGACTTGTGAAT 5843
184 yIleuIlyIleSerPheThrGlyIlyGlyIlySerGlyIlySerGlyIly 200
|||||
5844 ATCTAAAGATAGATTCTACGTAAGAAAGAAAGAAATATTCAAAAG 5893
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpIlyTrpGlyIly 217
|||||
5894 TGGATAAATGTATGAGCTGGGATAGTATTTTATTAATATGCGGGGG 5943
217 yAlaGlySerThrleuThrIleArgleuArglyIleGlyTrpGlyIly 234
|||||
5944 AGCAGGGTCACTTAAACATTCCTTGGATAGAGAGAGGAGGAGCAAC 5993
234 rObroValAlaMetGlyProAspIlyValIleuAlaGlyIlyGlyPro 250
|||||
5994 CCCCTGGGAGTGGGAGCCGATAAAGTACTGCTGAACAGGGGCCCGCC 6043
251 AlaIleuGlyProProHisAsnleuProValProglIleuThrSerleu 267
|||||
6044 GCCCTGAGACACCCGATTAATTCGCGGCCCCCAATTAACCTCGCTGCG 6093
267 gProAspIleuThrGlnProProSerAsnSerThrArglyleuIlePro 284
|||||
6094 GCCTGACATTAACACCCGCTAGCAGCGTACCACTGGATTGATTCCTTA 6143
284 hAsnThrProArgAsnSerProglYValProValIlyThrGlyIlyArg 300
|||||
6144 CCAACAGCGCTAGAAACCTCCAGAGTTCCTGTTAAGCAGAGAGAGA 6193
301 leuPheSerleuIleGlyIlyAlaPheGlyAlaIleAsnSerThrAsp 317
|||||
6194 CTCTTACGTCATCCAGGAGCTTCCCAAGCCATCACTCACGAGACC 6243
317 oAspAlaThrSerSerCysTrpIleuCysleuSerSerGlyProProTrp 334
|||||
6244 TGAATGCACTTCTTCTGTGTGCTTGTCTATCCACAGGGCTCTCTTAT 6293
334 yGluGlyMetAlaIlyGlyIlyArgIlyPheAsnValThrIlyGlyIlyArg 350
|||||
6294 ATGAGGGGATGGCTAAAGAGAAATTCATATGATCAACCAAGAGATAGA 6343
351 AsnGlyCysThrTrpGlySerArgAsnIlyleuThrIleuIlyValIle 367
|||||

```

6344 AATCATATGACATGGGGGTCGCCAATAAGCTTACCTCCTACTGAAAGTTTC 6393  
 367 TGLYSGLYTHRCYSILEGLYLSALAPROPROSERHISGLNHISLEUC 384  
 6394 CGGGAAGGAGCATGATAGGAAAGCTCCCAATCCACCAACACACTTT 6443  
 384 YSTYSER-THVALYALYRGUGINALASERGIUASNGILTYLLEUVAL 400  
 6444 GCTAATAGTACTGTGTTATAGCAGGCTCAGAAATCAGTATTATTACTA 6493  
 401 PROGLYTYRASNARGTRTPALACYSAN-THRGILEUTHRPROCYSA 417  
 6494 CTTGTTATTAACAGTGTGGGACATGCAATCTGCTTAACCCCTGTGT 6543  
 417 ISE-THSERVALPHEASNGILSERLYSASPLEUCYSVALMETVALGINI 434  
 6544 TTCCTCCHACGCTTCAACCAATCCAAAGATTCTGCTCATGGCCAAA 6593  
 434 LEVALPROARGVALTYRTHISPROGLUGILUVALLEUASPLGLTYR 450  
 6594 TCGTCCCGAGTACTACCTCCTGAGGAAGTGTCTCTGATGAAATAT 6643  
 451 ASPTYRATGYRASNARGPROLYSARGLUPROVALSERLEUTHRLEUAI 467  
 6644 GACTATCGGTATTAACCGACCAAAAGAGAACCCGATCCCTTACCTAGC 6693  
 467 AVALMETLEUGLYLEUGLYTHRALVALGLYVALGLYTHRGLYTHRALAA 484  
 6694 TGTAAATGCGCGATTAGGAGCGCGCTTGCGCTAGAACAGGACAGCTG 6743  
 484 LALEULIETHRGLYPROGLINGILUGLYGLYLEUGLYLLEUGLYLLEUHI 500  
 6744 CCGTATGATCAGAGCACAGCAGCTAGAGAAAGACTGGTGAAGCTACAT 6793  
 501 ALAALAMETTHRGILUASPLEUARGLALEULYSGILSERVALSERASLE 517  
 6794 GCGGCCATACAGAGATCTCCGACCTTAGAGAGTCTGTTAGCAACCT 6843  
 517 UGLIUGILSERLEUTHR-THSERLEUSERGLUVALLEUGILN-THSERARG 534  
 6844 AGAAGAGTCCCTGACTTCTTGTGAGAGTGTCTACAGAACCGAGG 6893  
 534 LYLEUASPLEU-THLEU-THLEU-THLEU-THLEU-THLEU-THLEU-TH 550  
 6894 GATTATATCTGCTGTTCTTAAAGAGAGGCTGATATGACGCTTAA 6943  
 551 GILUGILCYSCSPHETRYVALASPHIS-THSERGLYALILEARASPERME 567  
 6944 GAAGATGTTGCTTCTATATGATCACTAGAGACCATCAGACTCCAT 6993  
 567 TASNLYSLEUARGLYLYSLEUGILUARGARGARGARGILUARGILUARG 584  
 6994 GAGCAAGCTTAGAGAAAGGTAGAGAGGCTCGAAGGAGAAAGAGGCTG 7043  
 584 SPGLINGLYTRPHEGLUGLYTRPHEASNARG-THSER-THSER-THSER-TH 600  
 7044 ACCAGAGGCTGTTGAGAGATGTTCAACAGGCTCTCTGATGACACAC 7093  
 601 LEULEUSERVALLEUTHRGILPROLEUVALLEULEULEULEULEUTH 617  
 7094 CTGCTTCTGCTCTGAGGAGCCCTAGTAGTCTGCTCTCTTACTTAC 7143  
 617 VALGLYTR-THSERVALLEU-THSERVALLEU-THSERVALLEU-THSERVAL 634  
 7144 AGTTGGCCCTGCTTAAATTAATAGGTTGCTGCTTGTAGAGAACGAG 7193  
 634 ALSERVALVALGINILEMETVALLEUARGGLINGILTYLLEUGLYLEU 650  
 7194 TGGGTGACAGTCCAGATCATGATAGGACACAGTACCAAGGCTTCTG 7243  
 651 SERGLINGLYLTHRASPLEU 657  
 7244 AGCCAAAGAGAAAGTACCTC 7264

seq\_name: gp\_v1:PEN133818  
 seq\_documentation\_block: 8750 bp DNA VRL 14-Apr-2000  
 LOCUS: PEN133818 Porcine endogenous retrovirus type C proviral gag, pol and env  
 DEFINITION: genes and LTR (class B, clone 43).  
 ACCSSION: AJ133818  
 VERSION: AJ133818.1 GI:6688949  
 KEYWORDS: env gene; env protein; gag gene; gag protein; pol gene; pol protein.  
 SOURCE: porcine endogenous retrovirus.  
 ORGANISM: Porcine endogenous retrovirus.  
 REFERENCES: 1 (bases 1 to 8750)  
 1 Czauderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R. Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells  
 J. Virol. 74 (9), 4028-4038 (2000)  
 JOURNAL: 20219394  
 MEDLINE: 2 (bases 1 to 8750)  
 REFERENCE: Toenjes, R.R.  
 AUTHORS: Direct Submission  
 TITLE: Submitted (04-MAY-1999) Toenjes R.R., Medical Biotechnology, Paul-Enrich-Institut, Paul-Enrich-Strasse 51-59, Hessen, D-63225 Langen, GERMANY  
 FEATURES  
 source Location/Qualifiers  
 1..8750  
 /organism="porcine endogenous retrovirus"  
 /proviral  
 /strain="type C"  
 /db\_xref="taxon:61673"  
 /clone="43"  
 1..629  
 467  
 /note="cap site"  
 632..649  
 /note="gag"  
 1064..2635  
 /gene="gag"  
 1064..2635  
 /gene="gag"  
 /function="retroviral capsid"  
 /product="gag protein"  
 join(1064..2635,2639..6223,6099..8072)  
 /note="viral polypeptide"  
 /codon\_start=1  
 /protein\_id="CAB65341.1"  
 /db\_xref="GI:6688950"  
 /translation="MGQVYTPSLTLDHMTVRSRAHNLSTOVYKKGPMOTFCASEMP  
 TFDVGPSEGTNSSTILAVKALIFQVGGSSHPDDEPTLTMODLAEDPPWAKFWLN  
 KRKRDPRILALGEEKNHSKAEVSPSPRYPIEBPMPWEPQPPPPYPAQGVAVRG  
 PSAPDAPVVEAGTATRSRATPERDEIALILPRTGPPMGGLOPLDQAFPS  
 ADLYNKKTNHPPESDPDRLGLVESIMSHQPTMDCCOLLOTLETERERILLEA  
 RKVREGADSRPQLONEIDMORLPRTGPGDWYTAGRESIKTYRQALVAGASRRP  
 TNIAKREVMGPNPEPSVLETLMEARFRTPTDPTSAQASVALAFIGSALDIR  
 KKLQRLGLEAEALRDVLEAEKYYRRRETEKQKREKEREERDRQENLT  
 KILAAVEKSSSRERERDRKIRSGRSGNNGNTRPLDQCAVCKEGRHARNCPPK  
 KKKGKFKVLALEDKGRGSDPLEPRVALVEGQVPEFLVDTGAESHVLLQGLK  
 KKKSWMGATGQROVPMWTRTVDLAVGRVHSFLVIECPVPLIGRLDKMGAOI  
 SPFOGPREYVNNKPTTYVTLDDDEYRLXPOVKPDDIOSWLEOPPAWETAGMG  
 LAKOVPOVIOUKASATPVSRQYPLSRARGLIMPHVORLIOGGLVPVSPMPTPL  
 LPRKPGTIDRVDOLREVNRVODIHPTVNPYNLISALPERNWTVDLKAATF  
 CLRLHPTSQLPFAEWRDGTGTGOLITRLDQFKNSPTLEDAHLHDLANFRIQ  
 POYTLLOYDILLAGATKQDCIEGKALLILSLDGLYASAKKAOICREEVYLYGS  
 LGGGQWLEARKTYVOIIPATTAQVREFGTAGFCRLMIPGATLRAALPLYPKKE  
 KGFSAHPHOKTPAIRKALISAPALAPDVTKPTTYVDRKGVAGVLTOTLGPW  
 RRPVALISKKLPDASGWPCLKALIAVAALIVKDKDLTIGNITVYAPHAENIVRO  
 PDRMNTNARKMTHYOSLLTERVTEAPPAALPATLIDPETHEVTHDOKLIEETG  
 VRKDLIDPLTEGVLTWFTDSSYVEGKRMAGAAVVDGTHIMASSLPEGTSQAKAE  
 LMAITQALAEKGSINITYDSRYAFATNHAIGVATYKQGLTSGRLKNEEILTS  
 LEALHLPKRLATIIHCPGHOKARDLISRGNMADRVAKQAQAVNLLPIETPKADEPR

RQYTLDEMDIKKIDOFSEPEGTCTYSCKELLPHKESLEVVQOILRHLLTGLTKHLO  
 QLVATSPYHVLRLPGVADSVYVHCVPOLNANPSPGKRLRSGSHPGAHMVDTE  
 VKPAKYCNKYLIVDFPDSGMEVAPYPKETSTVYAKKILEELFPEPGIPKVGMS  
 PAFVAOVSQSLAKILGIDMKLHCAYRPOSSGOVARNKTKTELTTLTETGIDMDTA  
 LLEPVLEFRVNTPGGORTIPYELLXGSPRLYVIAASHADVLLSPLSRILKALBY  
 KORAMKOLREAYSEBDIYQVPHRQVODSYVYRHRNGLNLETWKGPYLVLLTTPYAV  
 KVEISTWIMHSHVKLAPPDSSGMRAREKTEENPKLRLHRLNYSNNNSGOMPLISM  
 RHLPTRGEPKRLRILPISFASIAMFLTLITPQASSKRLIDSNNHRLPLIWLITIDP  
 DTGVVNSTRGVAPRGTMWDELHCLRLINPAVSKSPNLVRSYGCPCPEKEKX  
 GGSSEFCRMSCVTSNDGMKWPISLODVEKSFVNSGPKYKWKLYKDKSCSPSD  
 LDYLTISFEKGOENIOMKINGMSGIVFYKGGAGSTLITRLITETEPYAVG  
 PDKYLABQGPALPEPHNLPVQOLTSRPIIDTPPSNGTGLIPYTPKNSPEVPKT  
 GQRFLSLIQAFQAINSTDPAITSCWLCISGPPYEGMAKGNVYKEHNOCTM  
 GSRNKLITLEVSGKTCIGKAPSHQHLCTVYVQASNOQYLPVGYNNMACNTGL  
 TPCVSSSVFNOSKDFCVMOIYPRVYHPEVYVDEYDRYRNPKEPPLTIAVWG  
 LGTAVGVTGTAALITGPOOLEKGLGELNAMEDLRALEESYNTESLTSISPYVL  
 ONRRGLDILFLREGGLCALKECCCFYVHSAIIRSMKESYRRLERRRRERADWG  
 FKGMFNRSMPMTJLTSALSLPLVLLLLITVGCPLINRFVAFVREKVSAYQIMVLKQO  
 YQGLSOGERTDL  
 2639..6223

## mat\_peptide

/gene="pol"  
 /function="retroviral protease and polymerase"  
 /product="pol protein"  
 2639..6223

## gene

## mat\_peptide

/gene="env"  
 /function="retroviral envelope"  
 /product="Env protein"  
 6099..8069

## gene

## LTR

## polyA\_signal

## BASE COUNT

2288 a 2214 c 2286 g 1962 t  
 ORIGIN

## alignment\_scores:

Quality: 3468.00 Length: 657  
 Ratio: 5.303 Gaps: 0  
 Percent Similarity: 99.543 Percent Identity: 98.021

## alignment\_block:

US-09-171-553B-10 x PEN133818 ..

Align seg 1/1 to: PEN133818 from: 1 to: 8750

1 Methisprothleusertrparginhisleuprothrarnglyglylupr 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6099 ATGCATCCACGTTAAGCTGGCGCACCTCCGACCTGGGGTGGAGAGCC 6148  
 17 olysargleuarqilleproleuserphealaserilealatrpheleut 34  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6149 GAAAGACTGAGAAATCCCTTAACCTTCGCTCCATCGCCCTGTTCTTA 6198  
 34 hrleuthrilethrproglinalaserislysarqleuileaspserser 50  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6199 CTCTAACAAATACCTCCAGGCCAGTAGTAAACCCCTTATAGACAGCTCG 6248  
 51 Asnprohisarqproleuserprothrtrpoulelleleasproasprh 67  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6249 AACCCTTAGACCTTATCTTACCTGGCTGATTAATGACCTCGATVAC 6298  
 67 rglyvalthrvalasnertharqglyvalalaproarglythtrp 84  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6299 GGGTGTACGTAAATAGCACTGAGGTGTTGCTCCTAGAGGACACTGGT 6348  
 84 rprogluleuhsphescysleuarqleuileasproalavalysser 100  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6349 GGCTGTAAGTCAATTCCTCCGATGATTAAACCCGCTGTTAAAGC 6398  
 101 Thrprohroasneleuarqsertryglyphetrycyscysprogllyth 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||

6399 AACCTCCCAACACTAGTCCGTAGTTATGGGTCTATGTGTCGCCAGGAC 6448  
 117 rglulysglulystrycysglyyserglygluserphecysarqargt 134  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6449 AGAATAAGAGAAATACCTGTGGGGTCTGAGGAATCTCTGTGAGAGAT 6498  
 134 rpsersvalthrserasnapolyasprtrpysrpproileserleu 150  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6499 GGACCTCGTCACCTCCACAGATGAGACATGGAATGCCGATCTCTCTC 6548  
 151 glinasprvallyspheserphevalasnerglyproglylystyryly 167  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6549 CAGAGCCGGGTAAATCTCTCTTGTCAATTCGGCCCGGCAAGTAAAC 6598  
 167 smelmellysleuthrlyasprlyssercysserproserasprleu 184  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6599 AGGATTAACACTATATTAAGATTAAGAGCTGCTCCCATCGACTTGATTT 6648  
 184 yrleulysileserphetrhrglulysglylysngluluanlileglnlys 200  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6649 ATCTAAAGATTAAGTTTCTACGAAAAAGGAAAAACAGGAAATATTCAAAAG 6698  
 201 trpilesnclymetsertrpglyllevalphetrylystyrylygl 217  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6699 TCGATTAATGTATAGCTGGGGAATAGTTTTTATTAATATGCGGGGG 6748  
 217 yalaglyserthrleuthrleuarqleuargileuuthrlythrglyclup 234  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6749 AGCAGGCTTCACCTTTTAACTTCGCTTAGAGTAGAGCGGGAGCAAGAC 6798  
 234 roprovalalamegliproasplysvalleuaglnllyprop 250  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6799 CCCCTGGCAGTAGGGAGCCGATTAAGTACTGCTGACACAGGGGCCCGG 6848  
 251 alaleugluproprohisasnleuprovalproglinleuthrserleu 267  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6849 GCCCTGGAGCCAGCGATTAACCTTCCGCGTCCCAATTAACCTCGCTGG 6898  
 267 gproasplethrghlproproserasnerthrtrglyleuileuprot 284  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6899 GCTGACATACACAGCCGCTAGCAACGGTACCTGATGATTCCTTA 6948  
 284 hrasnthrproarqasnerproglyvalprovalysthrlyglinarq 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6949 CCAACAGCGCTAGAACTCCCGAGGTCTTCGTTAAGACAGACAGACAGA 6998  
 301 leupheserleuileglnllyalapheginlialleasnerthraspr 317  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 6999 CTCTTACAGTCTATCCAGGAGGCTTTCACAGCCATCACTCCACCGACC 7048  
 317 oasppalathrsersercysrtrpouleuserserglypropotyrt 334  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7049 TGTATGCCACTTCTTCTTGTGGCTTTGTATTCCTCAGGGCTCTCTTAT 7098  
 334 yrglulymetalalysgluarqlysphasnvalthrlysglnhisarq 350  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7099 ATGAGGGGATGGCTAAGAGAAATTCATATGTGCCAAGACAGATAGA 7148  
 351 asnglincysrthrtrpilyserarqasnlyslleuthrleuthglvalise 367  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7149 AATCAATGTACATGGGGGCTCCGAATTAAGCTTACCTCAGGAAGTTTC 7198  
 367 rglylysglythrlycyslileglylvalalaproproserhisglnhisleuc 384  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7199 CGGGAAGGGGACATGATAGAAAACTCCCATCCCAACACACTTT 7248  
 384 ystyiserthrvalvaltryrglulinalasergluasnglntryleuval 400  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7249 GCTATAGTACTGAGTTATGAGCAAGGCTCAGAAAATCAGATATTAGTA 7298  
 401 proglytytyasnargtrpialacysasnhrnglyleuthrprocysva 417  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 7299 CCTGTTATACAGGAGGTGGGATGCAATACTGGTTAAACCCCTGTGT 7348



```

PDKVLAEOGPALPEPPHNLVPOLTSRLPDITQSPNSNGTGLIPTNTPRNSPGVEVKT
GDRFSLIGAFQALINSDPADTSSCMILCSGGPYEGMEAGKFNVTKEHRCCTW
GSRNKLTLFVSGKGTGKIGKAPPSHOHLCSTVYVYEOASENVYLVGYNRMACWTGL
TRCVSTSVSNQSKDRVMQVQIYPRVYIHEEVVIDEYDRYRPRRPSYSLFLAMTGL
LGTAVGVGTALITGPOLKEGJELHEAATFEDRALIEESVSNLESSTLSLEYVL
QNRNGIDLLFLREGGICAAKKECCFYVDHSGAIDBSMKLERLERRREREADOGW
FRGFWRSPMTTLLSALGPLVLLILLTVGPCLINRFVAVRERSVAVOIMVIRROO
YQILSOGSETDL"
mat_peptide
2729..6313
/gene="pol"
/function="retroviral protease and polymerase"
gene
2729..6313
/product="Pol protein"
/mat_peptide
6189..8162
/gene="pol"
/gene="env"
/function="retroviral envelope"
/product="Env protein"
gene
6189..8162
/gene="env"
LTR
8212..8918
polya_signal
8811..8816
BASE COUNT 2327 a 2245 c 2334 g 2012 t
ORIGIN
alignment_scores:
Quality: 3465.00 Length: 657
Ratio: 5.306 Gaps: 0
Percent Similarity: 99.391 Percent Identity: 98.021
alignment_block:
US-09-171-553B-10 x PEN133816 ..
Align seg 1/1 to: PEN133816 from: 1 to: 8918
1 MethisProthrLeuSerTrpArgHisLeuProThrArgGlyGlyLeu 17
:::|||||
6189 ATACATCCACGTTAAAGCTGGCGCACCTCCGACTCGGGSTGAGAGACC 6238
17 OlysarGleuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
6239 GAAAAGACTGAGAAATCCCTTAAGCTTCGCCCTCATCCCTGGTCTCTTA 6288
34 hrLeuThrIleThrProGlnAlaSerSerIleAspGluIleAspSerSer 50
|||||
6289 CTCTACAAATACCTCCAGGCCAGTACTAAAGCCTTATAGACAGCTCG 6338
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspThr 67
|||||
6339 AACCCCAATAGACCTTTATCCCTTACCTGGCTGATTTATGACCTGATAC 6388
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
6389 GGGTGTCACTGTAATAGCACTCGAGGTGTCTCCTAGAGGCACTGGT 6438
84 rPrProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
6439 GGCTGAAGCTGCATTTCTGCTCCGATTTGATTAAACCCGCTGTTAAAGC 6488
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyThr 117
|||||
6489 AACCTCCCAACCTAGCTGCTAGTATAGGTCTATTTGCTGCCAGGAC 6538
117 rGlyLeuGlyLeuSerGlyGlySerGlyGlyLeuSerPheCysArgArg 134
|||||
6539 AGAGAAAGAAATACTGTGGGATCTCTGGGAAATCCCTTCTGTAGGAGAT 6588
134 rPseCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeu 150
|||||
6589 GGAGCTGGCTACCTCCCAACGATGAGACTGAGAAATGGCGCATCTCTCTC 6638
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrpLys 167

```



```

7539 GACTATCGGTATACCGAACCAAAAGAGACCCGTATCCCTTACCCCTAGC 7588
467 aValmetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
7589 TGTATGCTCGGATTAGGACGCGCGCTGGCTAGAACAGCAGCAGCTCG 7638
484 lAlaLeuIleThrGlyProGlnGlnLeuGlyGlyLeuGlyGlyLeuLeu 500
7639 CCTGTATCACAGACACACAGCAGCTAGAAAGAGACTGTGTGCTACAT 7688
501 AlaAlaMetThrGluAspLeuArgAlaLeuGlyGlySerValSerAsnLe 517
7689 GGGCCCATGACAGAGATCTCCGACCTTAGAGAGTCTGTAGCAACCT 7738
517 uGluGluSerLeuThrSerLeuSerGlyValValLeuGlnAsnArgArg 534
7739 ACAAGAGTCCCTGACTTCTTGTGAGAGTGCTCTACAGAACCCGAGGG 7788
534 lLeuAspLeuLeuPheLeuArgGlyGlyLeuCysAlaAlaLeuLeu 550
7789 GATTGATCTGCTGTTCTTAGAGAGAGTGGGTATGTGCACCTTAAAA 7838
551 GluGluCysCysPheThrValAspHisSerGlyAlaIleArgAspSerMe 567
7839 GAAGATGTTGCTCTATGTAGATCATCTCAGAGCCATCAGACTCCAT 7888
567 tAsnLysLeuArgLysLysLeuGluArgArgArgGlyLysGlyLysAla 584
7889 GAGCAAGCTTAGAAAGGTTAGAGAGCGCTGGAAGGAAAGAGCGCTG 7938
584 sPglLysIleThrPheGluGlyThrPheAsnArgSerProThrMetThrTr 600
7939 ACCAGGGGTGCTTGAAGATGTTCAACAGSTCTCCTTGATGACACAC 7988
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
7989 CTGCTTCTGCTCTGACGGGACCCCTAGTAGCTCTGCTCTGTTACTTAC 8038
617 tValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlyLys 634
8039 AGTTGGCCCTGCTTAATTAAATAGGTTGTTGCTTGTAGAGAACGAG 8088
634 aLysAlaValGlnIleMetValLeuArgGlnGlnGlnGlnGlnGln 650
8089 TAGTGCATCTCCAGATCATGTAAGTCTAGCAACAGTACCAAGCCCTTCG 8138
651 SerGlnGlyGluThrAspLeu 657
8139 AGCCAAGAGAAACGTGACCTC 8159
seq_name: gb_pat:A66552
seq_documentation_block:
LOCUS A66552 8196 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 2 from Patent WO9740167.
ACCESSION A66552
VERSION A66552
KEYWORDS A66552.1 GI:4538105
SOURCE
ORGANISM
porcine endogenous retrovirus.
Virus; Retroviral; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE
1 (bases 1 to 8196)
AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.
TITLE PORCINE RETROVIRUS
JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;
O ONE BIOTECH LTD (GB)
FEATURES
location/Qualifiers
1..8196
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others
ORIGIN

```

```

alignment_scores:
Quality: 3445.50 Length: 658
Ratio: 5.276 Gaps: 2
Percent Similarity: 99.240 Percent Identity: 98.936
alignment_block:
US-09-171-553B-10 x A66552 ..
Align seg 1/1 to: A66552 from: 1 to: 8196
1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyLys 17
5606 ATGCATGCCACGTTAAGCGGGCCACTCCGAGCTCGGGGTGGAGAGCC 5655
17 oLysArgLeuArgLysLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34
5656 GAAAGAGCTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGCTGTTCTTA 5705
34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
5706 CTCTAACAAATACCTCCAGGAGGAGTAAAGCTTATAGACAGCTCG 5755
51 AsnProHisArgProLeuSerProThrTrpLeuIleLeuAspProAspTr 67
5756 AACCCCATACACCTTATACCTTACCTGCGTATATGACCTGATAC 5805
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
5806 GGGGTGCTACGTAAATAGCACTCAGAGGTGGTGCCTTAGAGGACCGGT 5855
84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
5856 GGCTGAACTGCAATTTCTGCTCCGATTTGATTAACCCCGCTGTTAAAGC 5905
101 ThrProAsnLeuValArgSerThrGlyPheThrCysCysProGlyThr 117
5906 ACACCTCCCAACCTAGTCCGATGATTAGGCTTATATGCTGCCCCAGCAC 5955
117 rGlyLysGlyLysThrCysGlyGlySerGlyGlySerPheLysArgArg 134
5956 AGAGAAAGAAAGAAATCTGTGGGGTCTGGGGAATCCTTCGTAGAGAGAT 6005
134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
6006 GGAGCTGCTACCTCCACAGATGAGACAGTGGAAATGGCCATCTCTCTC 6055
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysThr 167
6056 CAGACCGGGTAAATTCCTCTTGTCAATTCGCGCCGCGCAAGTACAA 6105
167 sMetMetLysLeuThrLysAspLysSerCysSerProSerAspLeuAsp 184
6106 AATGATGAACATATATAAGATTAAGAGCTGCTCCCATCAGACTAGATT 6155
184 rLeuLysIleSerPheThrGlyLysGlyLysGlnGlnLysIleGlnLys 200
6156 ATCTTAAGATTAAGTTTCACT...GAAGGAAACAGCAAAATATTCAAA 6202
200 sTrpLysGlnLysMetSerTrpGlyIleValPheThrLysThrGlyGly 217
6203 GTGGATAAATAGTATGAGCTGGGAAATGTTTATATTA...TATGCGCGGG 6251
217 lValGlySerThrLeuThrIleArgLeuArgIleGlyThrGlyThrGly 233
6252 GAGCAGGGTCCACTTAACATTCGCTTAGGATAGAGACGGGAGACAGAA 6301
234 ProProValAlaMetGlyProAspLysValLeuAlaGlnGlyPro 250
6302 CCCCTGTGGCAATGGACCGGATAAAGTACTGGCTGAAACGGGCGCCCC 6351
250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267

```

```

|||||
6352 GGGCCGGAGCCACCGCATTAACCTGGCGGCGCCCAATTAACTGCGTGC 6401
267 rtpProaspIleThrGlnProProSerAsnSerThrThyGlyLeuIlePro 283
6402 GGCTGACATTAACAGCGCGCTAGACACAGTACACAGTGGATTGATTCCT 6451
284 ThrAspThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300
6452 ACCACACCGCTAGAAACCTCCAGAGTGTCTCTTAAGACAGGACACAG 6501
300 gheupheserLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
6502 ACTCTTCACTTCATCCAGGAGCTTCCAGCCATCACTCCACCGAC 6551
317 rtsAspAlaThrSerSerCysThrLeuCysLeuSerSerGlyProProTy 333
6552 CTGATGCCACTCTCTCTGTTGGCTTGTCTATCTCTCAAGGGCTCTTAT 6601
334 TyrGlnGlyMetAlaLysGluArgLysPheAsnValThrLysGlnHisAr 350
6602 TATGAGGGGATGGCTAAAGAAATTCATATGACCAAGACATAG 6651
350 gsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlyValS 367
6652 AATCAATGTACATGGGGGTCCCGAAATAGCTTACCTCACTGAGATT 6701
367 ergLysGlyLysGlyLysAlaProProSerHisGlnHisLeu 383
6702 CCGGAGAGGGACATGATAGGAAAGCTCCCGCCACCAACACTT 6751
384 CysTyrSerThrValValTyrGlnGlnAlaSerGluAsnGlnTyrLeuVa 400
6752 TGCTATGACTGTGTATGATGACGAGGCTCAGAAATCAGTATTAGT 6801
400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417
6802 ACCTGTTATTAACAGGTGGTGGCATGCAATACGTGGTTAACCCCTGTG 6851
417 alserThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
6852 TTTCCACTCAGTCTTCAACCAATCCAAAGATTCTGTGTATGTGTC 6901
434 lIleValProArgValTyrTyrHisProGlnGluValValLeuAspGluTy 450
6902 ATCGTCCCGCAGTGTACTACCATCCGAGGAGGTGCTTATGATGATA 6951
450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467
6952 TGACATATGGTATTAACCGACCAAAAGAACCCGTATCCCTTACCTAG 7001
467 lValMetLeuGlyLeuGlyThrAlaValAlaGlyValGlyThrGlyThrAla 483
7002 CTGTATATCTCGGATTAGGAGCGCGCTGGCTAGGAGACGAGACAGCT 7051
484 AlAlaLeuIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlnLys 500
7052 GCCCGATCACAGGACACAGCAGCTAGGAAAGAGCTGTGTAGACTACA 7101
500 sAlAlaMetThrGlnAspLeuArgAlaLeuLysGluSerValSerAsnL 517
7102 TGGCGCCATGACAGAGATCTCCGAGCCCTTAAAGAGCTGTGTAGCAAC 7151
517 eugGlnGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
7152 TAGAAGATCCCTGACTTCTTGTCTGAGAGGTTCATACAGACCGAGAG 7201
534 GlyLeuAspLeuLeuPheLeuArgGlnGlyLysLeuCysAlaAlaLeuLys 550
7202 GGATTAGATCTGTGTTCTTAAGAGAAAGGTGCTTATGTGACGCTTAA 7251
550 sGlnGlnCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567
|||||

```

```

7252 AGAAGATGTTGCTTATGATGATCACTCAGGACCATAGAGACTCCA 7301
567 eAsnLysLeuArgLysLysLeuGluArgArgArgArgGluVala 583
7302 TGAACAGCTTAGAAAAAGTTAGAGAGGCGTGCAGAGGAAAGAGAGCT 7351
584 AspGlnGlyTrpPheGlnGlyTyrPheAsnArgSerProTrpMetThr 600
7352 GACGAGGGGTGTTGAAGAGATGTTCAACAGAGCTCTCTGGATGACAC 7401
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
7402 CTTCTTTCGTCTGTAGGGGCCCTAGTAGTCTGCTCTCTTACTTA 7451
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
7452 CAGTTGGGCTTGTCTTATTAATAGTTTGTGCTTGTGTAGAAACGA 7501
634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 650
7502 GTGAGTGCAGTCCAGATCATGTACTAGGCAACAGTACCAAGGCTTCT 7551
650 userGlnGlyGluThrAspLeu 657
7552 GAGCCAAAGAGAAACTGACTC 7573
seq_name: gb_v1:AF014162

```

```

seq_documentation_block: 1971 bp mRNA VRL 01-JUN-1998
LOCUS AF014162
DEFINITION Porcine endogenous retrovirus envelope protein (env) mRNA, complete cds.
ACCESSION AF014162
VERSION AF014162.1 GI:3169735
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1. (bases 1 to 1971)
AUTHORS
Haworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.
JOURNAL
2 (bases 1 to 1971)
TITLES
Haworth, C., Galbraith, D.N., Lees, G.M. and Smith, K.T.
Direct Submission
JOURNAL
Submitted (16-JUL-1997) Q-One Biotech Ltd., Todd Campus, Acre Rd., Glasgow G20 0XA, UK
FEATURES
source
1..1971
Location/Qualifiers
1..1971
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/cell_line="derived from cell line PK15"
/note="pOEV"
1..1971
/gene="env"
1..1971
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="AAC17948.1"
/db_xref="GI:3169736"
/translacion="MHPTLSRRLPTRGSPKRLRLPLSPASTAMFLTLITPOASSK
RLIDSNPHRLSLTWLILDPDGVYVNSRGVAPRGTMMDELHFLRLINPAVKSRP
PNLYRSYGFICCPETREKRYCGSGSEFRCRWSCVYSNDGKMPYLSLDNRVRFVN
SGPKRYMMKLYKDKSCSPSDLDYLTSTPRTKRYKSVKDYKELGNSFLVGGGAG
STLIRLRIETGTEPVAWGPDKVLAEGPPALPEPRLNLPDLITPDPSPS
TTLIPIPTNPSGVPYKVGORFLSLIOGAPFALINSTDPDSSCMCLSGSPYVE
GMAKERENYTKERNOCWTGSRMKLTLTFVSGKTCIGAPSHOILCYSTVYEOA
SENOYLTPGYNRMACNTGLTPCYSTSVNPODFCMYOITVPRVYTHNPEVYLDERD
YRINRKRREPVSLTAVMLGLGTRAVGVGTALITGPOOLEGLGELNAAKEDLKA
LKESVNSLESLSLSEVVLONRGDLTLRLRGGLCAALKECCFVYHDSGALRSM
NKLKKLERRRRERADGWFEGWFNRSPPMTTLLSLATGRLVLLLLLVGCLINR

```

BASE COUNT 530 a 490 c 482 g 469 t  
 FVAFRRERVSANQIMVLROQYGLLSQGETDL"  
 ORIGIN

## alignment\_scores:

Quality: 3444.50 Length: 658  
 Ratio: 5.275 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 98.936

## alignment\_block:

US-09-171-553b-10 x AF014162 ..

Align seg 1/1 to: AF014162 from: 1 to: 1971

```

1 MethIsProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
1 ATGCATCCACGTTAAGCCGGCCACCTCCGCACTCGGGGTGGAGAGCC 50
17 oluysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
51 GAAAGAGCTGAGATGCCCTTAAGCTTGCCCTCCATCGCTGTTCTTA 100
34 hIleuThrlIethrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
101 CTCTAACATTAACCTCCCGAGCCAGTAGTAAAGCCTTATAGACAGCTCG 150
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspH 67
151 AACCCCAATAGACCTTTATCCCTTACCTGGCTGATATTAGCCCTGATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
201 GGGTGTCACTGATAAATAGACACTCGAGGTGTGCTCCTAGAGGCACCTGT 250
84 rPProGlnLeuHISphecYsLeuArgLeuIleAsnProAlaValLysSer 100
251 GGGCTGAACGTGATTTGCTGCCATGATTAAACCCGCTGTAAAGC 300
101 ThrProAsnLeuValArgSerTrpGlyPheTrpYsCysProGlyTh 117
301 ACACCTCCCAACCTAGCTCGGTATGAGGTTCTATTGCTGCCCGAGGCAC 350
117 rGluLysGluLysTrpYsGlyGlySerGlyLysSerPhecYsArgArgT 134
351 AGAGAAAGAGAAATACTGTGGGGGTGTGGGAATCCCTCTAGAGAGAT 400
134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGCTCACCTCCCAAGATGAGACTGGAAATGGCCGATCTCTCTC 450
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
451 CAGGACGGGTAAATTTCTCTTGTGCAATTCGGGCCCGGCAAGTACAA 500
167 sMetLeuLysLeuTrpLysAspLysSerCysSerProSerAspLeuAspT 184
501 AATGATGAACATATATAAGATAGAGCTGCCCATCAGACTTAAGATT 550
184 yLeuLysLleSerPheThrGlyLysGlyLysGlnGluAsnIleGlnLys 200
551 ATCTAAGATTAAGTTTCACT...GAAAGGAAACAGGAAATATTCAAAA 597
200 strPileAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGlyLys 217
598 GTGGATTAATGGTATGAGCTGGGGAATACTTTTATTA.TATGGCGGGG 646
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyLysGlu 233
647 GAGCAGAGGTCCACTTTAACCATTCGCTTAGAGATAGAGCGGGAGACAGA 696
234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250

```

```

697 CCCCCTGTGGCAATGGACCCGATAAAGTACTGGCTGACAGAGGGCCCC 746
250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267
747 GGCCCTGGAGCCACCGCATTAACCTTGGCGTCCCAATTAACCTTCGGC 796
267 rGProAspLleThrGlnProProSerAsnSerTrpThrGlyLeuIlePro 283
797 GGCCGTGACATTAACACAGCGGCTAGCAACATACACTGATGATTATTCCT 846
284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300
847 ACCAACACGCTAGAAACTCCCGAGGTGTCTGTAAAGACAGAGACAG 896
300 gLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
897 ACTCTTCAGTCTCATCCAGGAGGCTTCCAAAGCATCAATCCACAGACC 946
317 rAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrp 333
947 CTGATGCCACTTCTTCTTGTGGCTTGTCTATCTCAGGGCCCTTAT 996
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350
997 TATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAG 1046
350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367
1047 AAATCAATGTACATGGGGGTCCGAAATTAAGCTTACCTCACTCAAGCTTT 1096
367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
1097 CCGGGAAGGGGACATAGCAATAGAAAGCTCCCACTCCACCAACACTT 1146
384 CysTrpSerThrValValTyrGluGlnAlaSerLysGlnGlyLysLeu 400
1147 TGTATATAGTACTGTGTATAGACAGGCTCCAGAAATACATGATTAGT 1196
400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCys 417
1197 ACCGTGTTATACAGGTGGTGGCATGCAATACTGGGTTAAACCCCTGTG 1246
417 aLserThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
1247 TTTCACCTCAGCTCTTCACCAATCCAAAGATTCTGTGCATGATGCCAA 1296
434 lLeValProArgValTyrTrpHisProGluGluValValLeuAspGlu 450
1297 ATCGTCCCGCAGTGTACTACCATCTGAGGAAGTGTCTTGTATGATA 1346
450 rAspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrLeu 467
1347 TGACTATCGGTATTAACGACCAAAAGAGAACCCGTTATCCCTTACCTAG 1396
467 lValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
1397 CTGTAATAGCTGAGTATGGAGCGGCGTGGCGTAGAGAACAGGACAGCT 1446
484 AlaleuIleThrGlyProGlnGlnLeuGluLysGlyLeuGluLeuHIS 500
1447 GCCCTGATCACAGGACACACAGCAGCTAGAGAAAGACTGTGTGAGCTACA 1496
500 sAlaAlaMetThrGlnAspLeuArgAlaLeuLysGlyLysSerValSerAsn 517
1497 TCGCGCCATGACAGAAAGATCTCCAGGCTTAAAGAGACTGTGTTAGCAAC 1546
517 euGlnGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
1547 TAGAAGAGTCCCTGACTTCTTGTGAAGGTGTTCACAGAACCGGAGG 1596
534 GlyLeuAspLeuLeuPheLeuArgGluGlyLysCysAlaLeuLys 550
1597 GGATTAGATGTGCTGTTCTAAGAGAGGTGGTTATGTGCAGCCTTAAA 1646

```

550 sgluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567  
 1647 AGAAGAAATGCTCTTCTATGATGACTCAAGAGACCTCAGAGACTCA 1696  
 567 eLaSnLysLeuArgLysLysLeuGluArgArgGluGluAla 583  
 1697 TGAACAACACTTGAAGAAAGTAGAGAGCGCTCGAAGGAAAGAGAGCT 1746  
 584 AspGlnGlyTrpPheGluGlyTyrPheAsnArgSerProTyrMetThr 600  
 1747 GACCAGGGGTGGTTGAAGAGTGTCAACAGGCTCTCTTGAAGACAC 1796  
 600 rLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 1797 CCGTCTTCTGCTCTGACGAGGCGCCCTGAGTACTCTCTCTGTTACT 1846  
 617 hValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633  
 1847 CAGTTGGGCGCTTCTTAATAGCTTGTGCTCTTGTAGAGACGA 1896  
 634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 650  
 1897 GTGAGTGCAGTCCAGATCATGCTAGGCAACAGTACCAAGGCTTCT 1946  
 650 uSerGlnGlyLeuThrAspLeu 657  
 1947 GAGCCAGAGAGAACTGACCTC 1968

seq\_name: gb.pat:A66553

seq\_documentation\_block:

LOCUS A66553 8209 bp DNA PAT 29-MAR-1999  
 DEFINITION Sequence 3 from Patent WO9740167.  
 ACCESSION A66553  
 VERSION A66553.1 GI:4538106  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM Porcine endogenous retrovirus  
 Viruses: Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 8209)  
 Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.  
 TITLE PORCINE RETROVIRUS  
 JOURNAL Patent: WO 9740167-A 3 30-OCT-1997;  
 Q ONE BIOTECH LTD (GB)

FEATURES  
 Source location/Qualifiers  
 1..8209

Source  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 588..2162

CDS

/product="VIRAL CORE PROTEIN"  
 /protein\_id="CAB39347.1"  
 /db\_xref="GI:4538107"  
 /translation="MGQVTPPLSLTLDHWEYRSRAHNLISVQKKGMQTCASEMP  
 TFDVGPSEGEENSEILAKIIFQTPGSHDQDEYILTMODLADPPWVKPMLN  
 KPRKGPRIALGEKNHSAKEVPESSYLPDRGADLAGTCTSPGLSTGCEG  
 TSAPGAPVVEPAGRSRSGATPERTDIALPRTGAPMPGGOLPLOWPES  
 ADLYNKTNNPRESDPORLTGLVESIMSHOETDCCOLOTLETTEREILLIA  
 RKVPGADGPTOLONEIDMGFLTPRGMDYNTARESEKIKYROLVAGLRASRP  
 TNLAKREVWQGNPEPVSFLERLMEAFRFTPTDTSEROKASVALATIGOSALDIR  
 KLLQRLGQLEALRDLVRAEKVYRRETEERKEDRKEERERERDRQEKMLT  
 KTLAAVEGKSSREREDFRKIRSGPROSGNLGNRTPLDKQCAVCYCKEGHARNCPK  
 KGNKGPVLALEDKXD"  
 <2163..5747  
 /codon\_start=1  
 /product="REVERSE TRANSCRIPTASE"  
 /protein\_id="CAB39348.1"  
 /db\_xref="GI:4538108"  
 /translation="GRGSDPLPEPRVTLKVEGQVPEPLVDGAEHSLLOPGLKSE  
 KSKMGATGQROYPTTTRTYDLGGRYTHSLVYPECPVPLGDLTKMAQSLKE  
 EQGPEVSNKPTIVTLTLDDEIRLYSPOVAPDDODIOSWLEQFOVAALETAGKLA

CDS

KQVPOVITOLKASATPVSVQYPLSREAREGIMPHVQRLIQOGLIVPQSPWMTPLP  
 VRKPTNDYRVPQDLREYNKRVQDILHPVPBPYMLLSALPERMWTYVLLDKAFEL  
 RLHPTNSQPLFAFEMRDPGTGRGOLUTWRLPQGFKNSTPIEDALHDLNFRLOHQ  
 VTILOYVDLILACATQDCEGTAKILLLEIDGYSASAKKAIQICREYTYLGSIR  
 GGQRLTTPARKKTYVQDPAPTTAKQYERLESTAFSCYMLIPGATITLAAPLYLTKEG  
 GFSVAPHEOKAFDITKALSLAPALADVDYKPTLTVDEKGVARKVLOTLPNR  
 PVAYLSKRLDVASGWPYCLKALIAVALLVNDAKILVGOITVIAHALENTVROP  
 DRMMINAMRTAYOSLLLEETFPAPALNPATLLPEETEPVPHDOLLEETGVK  
 KDLDIPLTGEVLTWETDSSVVEGKMACAAVVDGRTIMASLPEGTSAOKAEIM  
 ALTOALRIAEKSNINYDSRAFAATVHGAIVYKORGLTSAGREIKNKEILSILE  
 ALHLPKRLAITHCGHOKAKDLISRGNOMADRVAKQAQANVLLPIETPKAPPRRO  
 YLEDEWDEIKRIDPSETPGCTYSTGKELPRLBEGLEIVQOITRLITLTKHLOL  
 VRTSEYHVLRLPGVADSVYKPCQOLVNPASRIPEGKRLGSHPGAHWEVDEYK  
 PAKYGNKYLIVFVTFSGWVAEYPTKKTSTVYAKKILBELFPFRGIPKVIQSDNGPA  
 FVAQYOSGLAKILGIDMKLHCAYRPOSSGOYERNNRTIKETLTETGETGINDMALL  
 PEVLFRVNTPTGOGFLPTEYLIVGPGPLAIEAFASADYLTLSQPLSKALEWVR  
 RAMQOLRAYSGGDLQVPHRYQVDSVYRRHAGNLETFRKPKGYLVLLTPPAVKE  
 GIPTWHSYVKPAPPPDSGWRKATEMPLKRLRLHLPVTSNNNSPGO"  
 5620..7590  
 /codon\_start=1  
 /product="VIRAL ENVELOPE PROTEIN"  
 /protein\_id="CAB39349.1"  
 /db\_xref="GI:4538109"  
 /translation="MHPTLSRHLPTNGEPRKRLRIPSRASIMPLTLTTPPQASK  
 RLIDSSNPHRLSLTWLIIIDPDGTVAVSTGVAAPRGTMPELFLRLINPAVSTP  
 PNLVRSYFYCCPGTEKEREYCGGSESCVTRMNSCVTNDGDMKWPISLORVKSFVN  
 SGPGYKMMKLYKDKSCSPDLYLKISFTRKTKGSKYKVKVELNSFLYGGAG  
 STTIRLRIETGTEPVPAMGPDRLVLAEGPAPALPNHPOLTSRLPDITOPPS  
 TTGLIPMTNPNNSQVYKRGORLESLIGCAROAINSTDPDAPTSGLSSGPPE  
 GMAKRRKNTVKRHNOCCTWGSRRKNTLITVSGCTGICGAPRSHOILCTSYVEA  
 SENDLPVGINRWMACTGTLPCVSTSVFNOSKPCVAVQIVPRTVYHPEVILDEID  
 YRYNPKREPVSLTAVMLGIVAGVGTGIALITLPQOLEKGLHAAHMEEDLA  
 LKESVNLSESLTSLSEVILQNRGMLDLFLREGGLALAEKBCFYVDHSGALRDSM  
 NLRKKELERREREADQGEWGFENRSPWMTTLLSALTGLVLLVLLLTIVGPCILNR  
 FVAFYREVSVAVOIMVLRQYQGLISGETDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t

ORIGIN

alignment\_scores:

Quality: 3444.50 Length: 658  
 Ratio: 5.275 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment\_block:

US-09-171-553B-10 x A66553 ..

Align seg 1/1 to: A66553 from: 1 to: 8209

1 MethisProthLeuSerTrpArgHisLeuProthArgGlyGlyLupr 17  
 5620 ATGCATCCACGTTAAGCCGCGCCACCTCCACCTCGGGGTGAAGGCC 5669  
 17 oLysArgLeuArgLLeuProLeuSerPheHisSerIleAlaTrpPheLeu 34  
 5670 GAAAGACTGAGAACTCCCTTAAGCTCGCTCCATGAGCTTGGTTCCTTA 5719  
 34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSet 50  
 5720 CTCGTAACAATACATCCCGCAGGACAGTAGTAACGCTTATAGACAGCTCG 5769  
 51 AsnProHisArgProLeuSerProthTrpPheLeuIleIleAspProAspTh 67  
 5770 AACCCCATAGACCTTATACCTTACCTGCTGCTATATTGACCCCTGATAC 5819  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84  
 5820 GGGGTGCTAGTGAATAGACCTGAGGTGTGCTCCAGAGGACCTGGT 5869  
 84 rTrpGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 5870 GGCCTGAACTGCAATTTCTGCTCGCATTTGATTAAACCCCGCTGTTAAAGC 5919

```

101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyIth 117
|||||
5920 ACACCTCCCAACTAGTCCGATGATGCTTATGCTGCTCCAGGAC 5969
|||||
117 tglutglutylserCysGlyGlySerGlyGluSerPheCysArgArgT 134
|||||
5970 AGAGAAAGAGAAATCTGTGGGGGTCTGGGGAATCCTTCTGTAGAGAT 6019
|||||
134 tpserCysValThrSerAsnAspGlyAspTlpLysTTPProIleSerLeu 150
|||||
6020 GGAGCTGGCTACCTCCAAACGATGAGACTGGAATGCGCATCTCTCTC 6069
|||||
151 glnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
|||||
6070 CAGGACCGGTAATAATCTCTTGTCAATTCGCGCCCGGCAAGTACAA 6119
|||||
167 smetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
|||||
6120 AATGATGAACATATATAAGATAGAGCTGCTCCCATCAGACTAGATT 6169
|||||
184 tYrLeuLysIleSerPheThrGlyLys.GlyLysGlnGluAsnIleGln 200
|||||
6170 ATCTAAGATAAGTTTCACT...GAAGGAAACAGGAAATATTCAGAA 6216
|||||
200 strPileAsnGlyMetSerTTPGlyIleValPheTyrLysTyrGlyGly 217
|||||
6217 GTGGATAAATGCTATGAGCTGGGGAATAGTTTTTATTA.TATGGCGG 6265
|||||
217 lYalagLysThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233
|||||
6266 GAGCAGGGTCCACTTAAACATTCGCTTAGATAGAGACGGGGACAGAA 6315
|||||
234 ProProValAlaMetGlyProAspLysValLeuAlaGlnGlnLysPro 250
|||||
6316 CCCCCTGTGGCATGGAGCCGATAAAGTACTGGCTGAACAGGGCCCCC 6365
|||||
250 oAlaLeuGlnProProHisAsnLeuProValProGlnLeuThrSerLeu 267
|||||
6366 GGCCCTGGAGCCACCGCATACTTGGCGTGGCCCATTAACCTGGCTGC 6415
|||||
267 tgpProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283
|||||
6416 GGCTGACATACACAGCGCTACACAGATACACATGATGATGATCT 6465
|||||
284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGln 300
|||||
6466 ACCAACACGCTAGAAATCCCGAGGTCTCTGTTAAGACAGGACAGAG 6515
|||||
300 gLeuPheSerLeuIleGlnIleAlaPheGlnAlaIleAsnSerThrAsp 317
|||||
6516 ACTCTTCATCTCATCCAGGAGCTTCCAAAGCATCAACTCCAGCGAC 6565
|||||
317 roAspAlaThrSerSerCysTTPLeuCysLeuSerSerGlyProProTyr 333
|||||
6566 CTGATGCCACTCTCTTCTTGGCTTGTCTATCCTCAGCGCTCTCTAT 6615
|||||
334 TyrGlnGlyMetAlaLysGlnArgLysPheAsnValThrLysGlnHis 350
|||||
6616 TATGAGGGGATGGCTTAAGAAAGAAATTCATGTGACCAAGAGCATAG 6665
|||||
350 gAsnGlnCysThrTTPGlySerArgAsnLysLeuThrLeuThrGlnVal 367
|||||
6666 AATCAATGATCATGGGGGCCGGAATATAGCTTACCTCAGTGAAGTTT 6715
|||||
367 tGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
6716 CCGGGAAGGGGACATGATGAAAGCTCCCATCCACCAACACCTT 6765
|||||
384 CysTyrSerThrValValTyrGlnGlnAlaSerGlnAsnGlnTyrLeu 400
|||||
6766 TGCATATAGTACTGTGTTATGAGCAGGCTCAGAAATCAAGATTATAT 6815
|||||
400 lProGlyTyrAsnArgTTPTrpAlaCysAsnThrGlyLeuThrProCys 417
|||||
|||||
8816 ACCGTGTTATACAGGTGGGATGATCAATATCTGGTTAAACCCCTGTG 8865
|||||
417 alserThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
|||||
6866 TTTCCACCTTACGTCTTCAACCAATCCAAAGATTTCTGTGTCATGTCAA 6915
|||||
434 lIleValProArgValTyrTyrHisProGlnGlnValValLeuAspLys 450
|||||
6916 ATGCTCCCCGAGTACTACTCATCCGAGAGAGTGGTCTTATGATAA 6965
|||||
450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeu 467
|||||
6966 TGACTATCGGTATTAACCGACCAAAAGAGAACCCGATACCTTAACCTAG 7015
|||||
467 lalaMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyTyrAla 483
|||||
7016 CTGTAAATGCTCGGATTAAGGAGCGCTGGCGTAGGAAACAGGAGACGT 7065
|||||
484 AlaleuIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlyLysLeu 500
|||||
7066 GCCCTGATCACAGGACACAGACAGCTAGAGAAAGACTGTGAGCTACA 7115
|||||
500 salAlaMetThrGlnAspLeuArgAlaLeuLysGlnSerValSerAsn 517
|||||
7116 TCGCGCATGACAGAAAGATCTCCGAGCTTAAAGACTCTGTAGCAAC 7165
|||||
517 eugGlnSerLeuThrSerLeuSerGlnValValIleGlnAsnArgArg 533
|||||
7166 TAGAAGAGTCCCTGACTTCTTGTCTCAAGTGTCTTACAGAACCGAGG 7215
|||||
534 GlyLeuAspLeuLeuPheLeuArgGlyGlyLysCysAlaAlaLeu 550
|||||
7216 GGATTAGATCGCTGCTTCTTAAGAGAGGTGGTTATGTGACAGCTTAA 7265
|||||
550 sGlnGlnCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 567
|||||
7266 AGAAGAAATGTTGCTTATGATGATCACTCAGAGCATCATGAGACTCCA 7315
|||||
567 eLAsnLysLeuArgLysLysLeuGlnArgArgAlaArgGlnAla 583
|||||
7316 TGAACAAAGCTTAGAAAAAGTTAGAGAGCGCTCGAAGGAAAGAGAGCT 7365
|||||
584 AspGlnGlyTTPPheGlnGlyTTPPheAsnArgSerProTTPMetThr 600
|||||
7366 GACCAGGGGTGTTGAAGATGGTTCACACAGCTCTCTGATGAGACAC 7415
|||||
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
|||||
7416 CCTGCTTCTGCTCTGAGCGGGGCCCTAGATGCTCTCTCTCTGTTACTTA 7465
|||||
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgLys 633
|||||
7466 CAGTTGGGCTTCTCTTAAATTAATAGTTTGTGGCTTGTAGGAAGA 7515
|||||
634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeu 650
|||||
7516 GTGAGTGCAGTCCAGATCATGTACTTAGCCAAAGTACCAAGGCTTCT 7565
|||||
650 uSerGlnGlyLysThrAspLeu 657
|||||
7566 GAGCCAAAGGAGAACTGACCTC 7587
|||||
seq_name: gb_v1:PEN288589
seq_documentation block:
LOCUS PEN288589 1968 bp RNA VRL 29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B1 env gene for envelope
protein, genomic RNA.
ACCESSION AJ288589
VERSION AJ288589.1 GI:7414447
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.

```

## ORGANISM

porcine endogenous retrovirus  
 Viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.

## REFERENCE

1 (bases 1 to 1968)  
 Bosch, S., Arnaud, C. and Jestin, A.  
 Study of full-length porcine endogenous retrovirus genomes with  
 envelope gene polymorphism in a specific pathogen-free large white  
 swine herd

## JOURNAL

J. Virol. 74 (18), 8575-8581 (2000)

## MEDLINE

20411432

## AUTHORS

Bosch, S.

## TITLE

Direct Submission  
 Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,  
 AFSSA-Zoopole, Zoopole les Croix B.P.53, Ploufragan 22440, FRANCE

## FEATURES

Location/Qualifiers  
 1..1968  
 /organism="porcine endogenous retrovirus"  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:61673"  
 /clone="PERV-B1"  
 /note="French Large White SPF-pig"

## gene

## CDS

1..1968  
 /gene="env"  
 /codon\_start=1  
 /product="envelope protein"  
 /protein\_id="CAB86208.1"  
 /db\_xref="GI:7414448"  
 /translation="MPTLSMRQPTRGCEPKRLGIPLSFGSIAMFLVLTTPQSSK  
 RLIDSNPHNPSTLWLTIDPDGVVNSTKVAAPGTVWPELHCLRLINAVASTP  
 PNLVRSYGYCCPGTEKERYCGSGSEFCRMSVCVSNQDWMKWPISLDDRYKFSFN  
 SGKRYKVKLYKDKSCSPSDLYLKISFTEKGKQENIKKMGMSWGIYFKYGA  
 GSTLTITETGEPVAVGPKVLAEGPPALPPLPVLPPOLTSRDPDTPPSNGT  
 GLIPTPRSPGVYKTVKORLESLQGAFOAINSDPDATSSCWLCSSGPPVYEGM  
 AKGKFNVTKEHNOCTGSRNKLTITLTVSGGCTGCAKAPRHQHCYSTVYEGASE  
 NOYLVGVNRMACNTGLTPCVSTSTFENOSKPCVAVQVLPVRYTHPEMVIDEYR  
 YNRKREPVSLLIYVNLGLTAVGVTGAAXITGPQLEKLGELHMAETDLDALE  
 ESYNLSESLTSLSEVVLNRRGLDLLELREGLCALKEBCCFVVDHSGAIRDMSK  
 LRELRERRERADQGFENFENRSPMPTLLSLALGLVLLILLITVGPCLINRFV  
 AFVREVSVOIIMVLRMLKGEFOHTGGRY"

## BASE COUNT

522 a 483 c 492 g 468 t 3 others

## alignment\_scores:

Quality: 3340.50 Length: 645  
 Ratio: 5.252 Gaps: 1  
 Percent Similarity: 98.605 Percent Identity: 96.124

## alignment\_block:

US-09-171-553B-10 x PEN288589

Align seg 1/1 to: PEN288589 from: 1 to: 1968

```

1 MethSProthLeuSerTTPATGHisLeuProthArgIlyGlyIupr 17
  |||||||
1 ATGATCCACAGTTAAGCTGCGCCAAACCCGACCTCGGGGTGGGAGCC 50
  |||||||
17 oLyArgLeuArgIleProLeuSerPheAlaSerIleAlaIleProPheLeu 34
  |||||||
51 GAAAGACTGGGAATCCCTTAAGCTTGGCTCATCGCTGGTCTCTTA 100
  |||||||
34 hrLeuThrIleThrProGlnAlaSerSerIlyArgLeuIleAspSerSer 50
  |||||||
101 CTCTNACAAATTAATCCCGAGCGAGTAGTAAGCCCTTATAGACAGCTCG 150
  |||||||
51 AsnProHisArgProLeuSerProthThrProLeuIleIleAspProAspTh 67
  |||||||
151 AACCCCATAGACCTTATCCCTTACTGCTGATTAATGACCCGTATAC 200
  |||||||
67 rGlyValThrValAsnSerThrArgGlyValAlaIleProArgGlyThrTrp 84
  |||||||

```

```

|||||
201 GGGGTGCTACTGTAAATAGACACTGGGGGTGTGGCTTCCTAGAGGACCTGGT 250
  |||||||
84 rPrProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValIlySer 100
  |||||||
251 GGCTTGACATGCATTTCTGCTCGATGATTAACCCCGCTGTAAAGC 300
  |||||||
101 ThrProAsnLeuValArgSerThrGlyPheThrCysCysProGlyThr 117
  |||||||
301 ACACCTCCCAACCTAGTCCGTAGTATGGTTCTATTGCTGCCAGGAC 350
  |||||||
117 rGlyLysGlnLysThrCysGlyLysSerGlyIleSerPheCysArgArgT 134
  |||||||
351 AGAAGAAAGAGAAATACTGTGGGCTTCTGGGGAATCCTTGTAGAGAT 400
  |||||||
134 rPserCysValThrSerAsnAspGlyAspTrpIlyStrProIleSerLeu 150
  |||||||
401 GGAGCTCGTCCACCTCCAGCAGATGAGACGAAATGGCCGATCTCTC 450
  |||||||
151 GlnAspArgValIlyPheSerPheValAsnSerGlyProGlyLysThrIly 167
  |||||||
451 CAGACCGGGTAAATTTCTCTTTGTCAATTCGGGCCGGGCAAGTACAA 500
  |||||||
167 sMetMetLysLeuThrLysAspLysSerCysSerProSerAspLeuAspT 184
  |||||||
501 AGTGATGAACCTATATTAAGATTAAGAGCTGCTCCCATCAGACTTAGATT 550
  |||||||
184 rIleLysLysIleSerPheThrGlnLysGlyLysGlnLysIleGlnLys 200
  |||||||
551 ATCTAAAGATTAAGTTTCACGAAAGAAAGAAAGAAATATATCAAAAG 600
  |||||||
201 TrpIleAsnGlyMetSerTrpGlyIleValPheThrLysThrGlyLysI 217
  |||||||
601 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
  |||||||
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGlyThrGlyThrGln 234
  |||||||
651 AGAGAGGTCCACTTAACCATT.....ATAGAGAGGGGGACAGAAC 691
  |||||||
234 rProValAlaMetGlyProAspLysValIleuAlaGlnGlyPropo 250
  |||||||
692 CCCCTGTGGAGTGGAGACCCGATTAAGTCTGCTTAACAGGGGCCCCCG 741
  |||||||
251 AlaLeuGlnProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
  |||||||
742 GCCCTGGAGCAGCCGATTAAGTTCGCCGTGCCCATTAACCTCGGTGG 791
  |||||||
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 284
  |||||||
792 GCCTGACATTAACACACCCGCTAGAACGCTACACTGATTAATCCCTA 841
  |||||||
284 hrAsnThrProArgAsnSerProGlyValProValThrGlnGlnArg 300
  |||||||
842 CCAACACGCTTAGACACTCCCGAGGTGTTCTGTTTAAGCAGACAGAGA 891
  |||||||
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317
  |||||||
892 CTCCTTAGTCTCATCCAGGAGCTTCCAAAGCATTCACTCAGCAGACCC 941
  |||||||
317 oAspAlaThrSerSerCysTrpLeuLysLeuSerSerGlyPropoIlyT 334
  |||||||
942 TATATGCACTTCTTCTTGTGGCTTGTCTATCTACCTAGGCTCTCTATT 991
  |||||||
334 yGlnGlyMetAlaLysGlnArgLysPheAsnValThrLysGlnHisArg 350
  |||||||
992 ATGAGGGGATGGCTTAAGAGAGAAATTAATGTAACCAAGAGCATGAGA 1041
  |||||||
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlyValSe 367
  |||||||
1042 AATCATATGATGAGGGGGGTCGGAATTAAGCTTACCTCAGCAAGATTTC 1091
  |||||||
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384
  |||||||

```

1092 CGGAGAGGAGCATGATAGAGAAAGCTCCCATCCACACACCTTT 1141  
 384 ySYrSerThvAlValIYrGlInAlaSerGluAsnGlnTyrLeuVal 400  
 1142 GCTATAGACTGTANTTTATGAGCAGGCGCTCAGAAATCACTATTAGTA 1191  
 401 ProGlyTyrAsnArgTPTPrAlaCysAsnThrGlyLeuThrProCysVa 417  
 1192 CCGTGTATTAACAGGTGGTGGCATGCAATACTGGTTAACCCCTGTCT 1241  
 417 IserThSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434  
 1242 TTCCACCTCAGCTTCAACCAATCCAAAGATTCTGTGTATGGTCCAAA 1291  
 434 IValProArgValIYrTyrHisProGluGluValValLeuAspGlnTyr 450  
 1292 TCGTCCCGGAGGTACTACATCCTGAGAAATGGTCTTGATGAAATAT 1341  
 451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467  
 1342 GATATTCGGTATTAACCGACCAAAAAGAGAACCCGATCCCTAACCATAGC 1391  
 467 aValMetLeuGlyLeuGlyThrAlaValAlGlyValGlyThrAlaAla 484  
 1392 TGTATAGCTCGGATTAGGAGCGCGCTTGCGGTAGAGAACGAGCAGCTG 1441  
 484 IAlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyLeuGlyGlnLeuHis 500  
 1442 CCGTATACACAGACACACAGCAGTTAGAGAAAGACCTTGAGAGCTACAT 1491  
 501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 517  
 1492 GCGGCCATGACAGAAAGATCTCGAGCCTTAGAGAGTCTGTAGAACACT 1541  
 517 uGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534  
 1542 AAGAGAGTCCCTGACTCTTGTCTGAAGTGGTCTACAGAACCGAGAGG 1591  
 534 ILeuAspLeuLeuPheLeuArgGluGlyGlyLeuGlyAlaAlaLeuLys 550  
 1592 GATTAGATCTGCTGTTCTAAGAGAGGTGGGTGTGTGACGCTTAAAA 1641  
 551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567  
 1642 GAAGAGATGTCTCTATATAGATCACCAGAGCATTACAGAGATCCAT 1691  
 567 fAsnLysLeuArgLysLysLeuGluArgArgArgGluAlaAla 584  
 1692 GAGCAAGCTTGAAGAAAGGTAGAGAGCGTCTGAGAGGAAAGAGCGCTG 1741  
 584 sGlnGlyTyrPheGlnGlyTyrPheAsnArgSerProTyrMetThrThr 600  
 1742 ACCAGGGGTGTTGAGAGATGTTCAACAGGTCTCTGATGATGACCCAC 1791  
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 1792 CTGCTTCTGCTCTGACGGGACCCCTAGTACCTGCTCTGTACTTAC 1841  
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 634  
 1842 AGTTGGGCTTGTCTTAATTAAATAGTTTGTCTTGTAGAGAACGAG 1891  
 634 alSerAlaValAlaGlnIleMetValLeuArgGlnGln 645  
 1892 TGAGTCACTCCAGATCATGTACTTAGAACACGC 1926  
 seq\_name: gb\_v1: PEN288588  
 seq\_documentation\_block: 2000 bp RNA VRL 29-AUG-2000  
 LOCUS PEN288588 2000 bp RNA VRL 29-AUG-2000  
 DEFINITION Porcine endogenous retrovirus PERV-AB7 env gene for envelope  
 protein, genomic RNA.  
 ACCESSION AJ288588

VERSION AJ288588.1 GI:7414445  
 KEYWORDS env gene; envelope protein.  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus  
 Viruses; Retroid viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 REFERENCE 1 (bases 1 to 2000)  
 AUTHORS Bosch, S., Arnaud, C. and Jestin, A.  
 TITLE Study of full-length porcine endogenous retrovirus genomes with  
 envelope gene polymorphism in a specific-pathogen-free large white  
 swine herd  
 JOURNAL J. Virol. 74 (18), 8575-8581 (2000)  
 MEDLINE 20411432  
 REFERENCE 2 (bases 1 to 2000)  
 AUTHORS Bosch, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,  
 AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE  
 FEATURES  
 source  
 1..2000  
 /organism="porcine endogenous retrovirus"  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:61673"  
 /clone="Perv-AB7"  
 /note="French Large White SPF-pig"  
 1..606  
 /gene="env"  
 1..606  
 /note="truncated, not functional"  
 /codon\_start=1  
 /product="envelope protein"  
 /protein\_id="CA86207.1"  
 /db\_xref="GI:7414446"  
 /translation="MHPTLSMRHLPTREGPEPKRLRIPSPASIAMFLUTLTPQASSK  
 RLIDSSNHRPLSTLWLIDPDGTGVSTGVAARGVWPELHRCRLIPAVKSTP  
 PLVLSYFPCPGTEKRYCGSGSESCRKRSCTSDGIMKPIISLDKVKSFVN  
 SGPKRYKWKLYKDKSCSPSDLDYLFKISFTKGNRKIPKSG"  
 BASE COUNT 532 a 496 c 499 g 473 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 3340.00 Length: 646  
 Ratio: 5.219 Gaps: 1  
 Percent Similarity: 99.071 Percent Identity: 96.749  
 alignment\_block:  
 US-09-171-553B-10 x PEN288588 ..  
 Align seg 1/1 to: PEN288588 from: 1 to: 2000  
 1 MethisProThrLeuSerTyrPArgHisLeuProThrArgGlyGlyIupr 17  
 1 ATGCATCCACGTTAAAGCTGGCGCCACCTCCGACCTGGGGGTGAGAGCC 50  
 17 olysArgLeuArgIleProLeuSerPheAlaSerIleAlaIleProPheLeu 34  
 51 GAAAGGCTGAGAAATCCCTTAAGCTTGCCTTCATCGCTGAGCTGCTCTTA 100  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 101 CTCTAACCAATTAACCTCCAGGCGAGTGAATAAACCTTTATAGACAGCTG 150  
 51 AsnProHisArgProLeuSerProThrTyrPheLeuIleLeasProAspTh 67  
 151 AACCCCATAGACCTTTATCCCTTACCGGCTGATTTAGACCTGATAC 200  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTyr 84  
 201 GGGTGTCACTGTAATAATAGACTGAGGTGTGCTCTTGAAGACGACCTGT 250  
 84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

```

|||||
251 GGCCTAACTGCTTTCTGCTCCGATGATTACCCGCTGTAAAC 300
|||||
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
301 AACACTCCCACTACTCCGCTAGTATAGGTTCTATGCTGCGCAGAC 350
|||||
117 rGluGlySerGlyTyrCysGlyGlySerGlyLysSerPheCysArgArg 134
|||||
351 AAGAGAGAGAAATACTGTGGGGTCTGGAGAACTCTCTGTAGAGAT 400
|||||
134 rPseCysValThrSerAsnAspGlyAspTrpLysTrpProLysLeu 150
|||||
401 GGAGCTGGCTCACTCCCAACGATGAGACTGAGAAATGGCCGCTCTCTC 450
|||||
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyr 167
|||||
451 CAGGACCGGGTAAATTCCTCTTGTCAATTCGCGCGGGAAGTACAA 500
|||||
167 smetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184
|||||
501 AGTGATGAACCTATATAAAGATAGAGCTGCCCATCAGACTTAGATT 550
|||||
184 yLeuLysLLeSerPheThrGluLysGlyLysGluAsnLLeGlnLys 200
|||||
551 ATCTAAGATAGTTTCACTGAAAGAGAA. CAGGAAATATTTCAGAG 599
|||||
201 TrpIleAsnGlyMetSerTyrPglyLLeValPheTyrLysTyrGly 217
|||||
600 TGGATTAATGTATAGCTGGGGAATAGTTTATTAATATGGGGGG 649
|||||
217 yAlaGlySerThrLeuThrIleArgLeuArgLLeGluThrGlyThrGlu 234
|||||
650 AACAGGCTCACTTAACTTACCATTCGCTTGAATAGAGACGGGACAG 699
|||||
234 rOProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 250
|||||
700 CCCCTGTGCGAGTGGGAGCCGATTAAGTACTGGCTGAACAGGGCCCC 749
|||||
251 AlAlaGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267
|||||
750 GCCCTGGAGACCCGATTAACCTCCGCTGCCCAATTAACCTCGTCG 799
|||||
267 gProAspLLeThrGlnProProSerAsnSerThrThrGlyLeuLeu 284
|||||
800 GCTGACTTAACACAGCCGCTACCAACGATGAGTGAATGATCTTA 849
|||||
284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300
|||||
850 CCAACAGCCCTAGAAACTCCCGAGTGTCTGTATAGACAGACAGAGA 899
|||||
301 LeuPheSerLeuLLeGln. GlyAlaPheGlnAlaLLeAsnSerThr 317
|||||
900 ATCTTCAGTCTCATCCAGGGAGCTTCCAAAGCCCACTCCACAGCAC 949
|||||
317 rAspAlaThrSerSerCysTrpLeuGlyLysSerSerGlyProPro 333
|||||
950 CTGATGCCACTCTCTTGTGTGTGTCTATCTCAGGGCCCTCTAT 999
|||||
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlu 350
|||||
1000 TATGAGGGAGTGGCTAAAGAGAAATTTAAAGTGACCAAGAGACCTAG 1049
|||||
350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGlu 367
|||||
1050 AAATCAATGTACATGGGGGTCCGAAATAGTTACCTCAGCAAAATT 1099
|||||
367 erGlyLysGlyThrCysLLeGlyLysAlaProProSerHisGlnLys 383
|||||
1100 CCGGGAAGGGGACATGATAGAAAGCTCCCATCCACCAACACCTT 1149
|||||
384 CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlyThrLeu 400
|||||
|||||
1150 TGTATAGTACTGTGTTATGACGAGCCCTCAGAAATAGATATTAGT 1199
|||||
400 LProGlyTyrTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrPro 417
|||||
1200 ACTGGTTGTACAGAGGTGGGATGATCAATACGAGGTTAAACCCCTG 1249
|||||
417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
|||||
1250 TTTCACCTAGTCTTCAACCAATCCAAAGATTCTGTGTCATGTCCTCA 1299
|||||
434 IleValProArgValTyrTyrHisProGluGluValValLeuAsp 450
|||||
1300 ATCTGCCCGAGTGTACTACCATCCGAGGAAGTGTCTCTGATGATTA 1349
|||||
450 rAspTyrArgTyrAsnArgProLysArgGlyLysProValSerLeuThr 467
|||||
1350 TGACTATCGGTATACCGACCAAAAGAACCCGATTCCTTACCTAGCT 1399
|||||
467 lAlaMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
1400 CTGTATATGCTCGGATTTAGGAGCGCGTGGCTAGGAACAGGACAGCT 1449
|||||
484 AlAlaLeuLLeThrGlyProGlnGlnLeuGluLysGlyLeuGlyL 500
|||||
1450 GCCCTGATCACAGACACACAGCCTAGAGAAAGACTGTGTGACCTACA 1499
|||||
500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGlyLysValSerAsn 517
|||||
1500 TGGCGTATATACAGAAAGATTCGAGACCTTAGAGAGTGTGTTAGCAAC 1549
|||||
517 euGluGlnSerLeuThrSerLeuSerGlyValValLeuGlnAsnArg 533
|||||
1550 TAGAAGAGTCCCTGACTTCTTGTCTGAAGTGTCTACAGAAACGGAG 1599
|||||
534 GlyLeuAspLeuLeuPheLeuArgGlyGlyGlyLeuCysAlaAla 550
|||||
1600 GGAATTAATGCTGCTGTTTAAAGAGAGTGGGTTATGTGACGCTTAA 1649
|||||
550 sGluGluCysCysPheTyrValAspHisSerGlyAlaLLeArgAsp 567
|||||
1650 AGAAGATGTGCTTCTATGTATGATCATTACAGAGCATTACAGACTCCA 1699
|||||
567 etAsnLysLeuArgLysLysLeuGluArgArgArgLysGluArgL 583
|||||
1700 TGAGCAAGCTTAGAGAAAGTTAGAGAGCGTCGAAGGAAAGAGAGCT 1749
|||||
584 AspGlnGlyTrpPheGlnGlyLysTrpPheAsnArgSerProTrp 600
|||||
1750 GACCAAGGGTGTGTTGAAGATGTTCACACAGTCTCTTGGATGACAC 1799
|||||
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
|||||
1800 CCGTCTTGTGCTGTGACGGAGCCCTAGTACTCTGCTCTGTTACTTA 1849
|||||
617 hrValGlyProCysLeuLLeAsnArgPheValAlaPheValArgLys 633
|||||
1850 CAGTTGGGCTTGTCTTAATTAATAGTGTGCTGCTTTAGAGAAAGA 1899
|||||
634 ValSerAlaValGlnLLeMetValLeuArgGlnGln 645
|||||
1900 GTGAGTGCAGTCCAGATCATGTGTTAGTACAAACG 1935
|||||
seq_name: gb_v1:PEN288592
seq_documentation_block:
LOCUS PEN288592 2000 bp RNA VRL 29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B7 env gene for envelope
protein, genomic RNA.
ACCESSION AJ288592
VERSION AJ288592.1 GI:7414453
KEYWORDS env gene; envelope protein.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus

```



REFERENCE  
 AUTHORS  
 TITLE  
 1  
 Viruses; Retrovird viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 (bases 1 to 2000)  
 Bosch, S., Arnauld, C. and Jestin, A.  
 Study of full-length porcine endogenous retrovirus genomes with  
 envelope gene polymorphism in a specific-pathogen-free large white  
 swine herd  
 J. Virol. 74 (18), 8575-8581 (2000)  
 20411432  
 JOURNAL  
 MEDLINE  
 2 (bases 1 to 2000)  
 Bosch, S., Arnauld, C. and Jestin, A.  
 Study of full-length porcine endogenous retroviral (PERV) genomes  
 with envelope gene polymorphism in a Large White  
 specific-pathogen-free swine herd  
 Unpublished  
 3 (bases 1 to 2000)  
 Bosch, S.  
 Direct Submission  
 Submitted (28-MAR-2000) Bosch, S., Molecular Biology Department,  
 AFSSA-Zoopole, Zoopole Les Croix B.P. 53, Ploufragan 22440, FRANCE  
 JOURNAL  
 TITLE  
 Location/Qualifiers  
 1..2000  
 /organism="Porcine endogenous retrovirus"  
 /specific\_host="Sus scrofa"  
 /db\_xref="taxon:61673"  
 /clone="PERV-B7"  
 /note="French Large White SPP-plg"  
 1..1899  
 /gene="env"  
 1..1899  
 /gene="env"  
 /codon\_start=1  
 /product="envelope protein"  
 /protein\_id="CAB86211.1"  
 /db\_xref="GI:7414454"  
 /translation="MPTLSMRHLPTRGERRKRLIPSPASFAMFLTLITIIPOASK  
 RLIDSNPHRLSLIMLITDPDVTNSTRGVARGTMMPLHCLRLINAVASTP  
 PNLVRSYGCPCPTKEKKEKCGSGESRCRRCVTSNDGDKMPILFODRKFSFN  
 SGGKTRKMLIKDKSCSPSDIDYKISTEKGKOBNIQKJNGSMGIVFKYGGGA  
 GSTVILRLRIETGTEPPVAVGPDVLAEGPAPLPPHLPVPLTSLPDTIOPPSN  
 GTTGLPTMPNPNBPVKTGKRLKLSLIGQAFKIKSTDPAANSQWCLSSGPY  
 EGMKEGKNVTKRHNQCTGSRNKLTLVSGKGTGIGKAPSHOHPCSTVVEQ  
 ASENQVLYGVNMMACNTGLTPEVSTSVENSGKPCVAVIIPRYVYHPEVDEY  
 DYRNRPKGEVPTITLAVMLGIGTAVGVTGTAALITGPQLEKIGELHAAMTEDLR  
 ALPESTSNLESESTISLSSVLDNRGIDILRLREGICAALEKECCFYDHSGLRDS  
 MSKRLRLRRRREREDQGFEGMFRNSPMTTLTSLATGLVYLVLITLITVPCLIN  
 RFVAFVRE"  
 BASE COUNT 532 a 485 c 505 g 477 t 1 others  
 ORIGIN  
 alignment\_scores:  
 Quality: 3329.00 Length: 645  
 Ratio: 5.243 Gaps: 0  
 Percent Similarity: 98.450 Percent Identity: 95.814  
 alignment\_block:  
 US-09-171-553b-10 x PEN288592 ..  
 Align seg 1/1 to: PEN288592 from: 1 to: 2000  
 1 MethiSPROTHLeuSerTTPArgHisLeuProThrArgGlyGlyLupr 17  
 1 ATGCATCCACGCTTAAGCTGAGCCACCTCCGACTCGGGGTGAGAGCG 50  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 1 GAAAGACTGAGAAATCCCTTAAGCTTCGCTGCTGCTGCTGCTTA 100  
 34 hLeuThrIleThrProGlnAlaSerSerIleArgLeuIleAspSerSer 50  
 101 CTCTAACAAATATCCCGACGACGAGTAAAGCCCTTATGACAGACTCG 150

51 AsnProHisArgProLeuSerProThrTrpLeuIleAlaAspProAspPth 67  
 151 AACCCCATAGACCTTATCCCTTACCTGCGGATTTATGACCCGTGATAC 200  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrPPT 84  
 201 GGGTGCTACTGTAATAAGACACGAGGTGCTGCTGCTGAGGACCTGTGT 250  
 84 rProGlnLeuHisPheCysIleArgLeuIleAsnProAlaValLysSer 100  
 251 GGCTGTAATGCTATTCGCTCCGATGATTAACCCCGCTTTAAAGC 300  
 101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyYth 117  
 301 ACACCTCCCAACCTAGTCCGTATGATGCTTCTATGCTGCCAGGAC 350  
 117 rGlyLysGlyLysTrpCysGlySerGlyLysSerGlyLysSerPheCysArgArgT 134  
 351 AGAGAAAGAGAAATACGTGGGGGTCTGGGGAATCCTTCTAGAGAT 400  
 134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeu 150  
 401 GAGAGCTGCTACCTCCACGATGAGACTGCAATGCGCATCTTCTTC 450  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrpLys 167  
 451 CAGGACCGGGTAAATTCCTTCTTGTCAATTCGGCGCCGGAAGTACAG 500  
 167 sMetLeuLysLeuTrpLysAspLysSerCysSerProSerAspLeuAspT 184  
 501 AGTGATGAAACTATATTAAGATTAAGAGCTGCTCCCATAGACTAGATT 550  
 184 YLeuLeuLysLeuSerPheThrGlnLysGlyLysGlnLysLysLysLys 200  
 551 ATCTAAGATTAATTTCTACTGAAAAAGAAACAGAAATATTTCAAG 600  
 201 TrpIleAsnGlyMetSerTrpGlyLysValPheTrpLysTrpGlyLys 217  
 601 TGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650  
 217 YAlaGlySerThrLeuThrIleArgLeuArgGlyLysTrpGlyLys 234  
 651 AGCAGGCTCCACGTAACTTCCTTGGATGATGATGATGATGATGATGAT 700  
 234 rProValAlaMetGlyProAspLysValLeuAlaGlnLysLysLysLys 250  
 701 CCCTGCTGAGGAGTGGGACCCGATTAAGTACTGCTGACAGGAGGCGCC 750  
 251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267  
 751 GCCCTGAGGACCGGCACTTCCGATGCTGCTGCTGCTGCTGCTGCTG 800  
 267 gProAspIleThrGlnProProSerAsnSerTrpThrGlyLeuIleProT 284  
 801 GCCCTGATTAACACAGCCGCTTACCAAGGTTACCATGCTGATTTCTTA 850  
 284 hAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300  
 851 CCAACATGCTAGAACTCCCGAGGTGCTCTGTTAAGACAGAGAAAGAGA 900  
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspP 317  
 901 CTCTCAGTCTCATCCAGGAGCTTTTAAAGCATCAAAATCCACGCGCC 950  
 317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTrpT 334  
 951 GGAGGCCAATTTCTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1000  
 334 YrGlyLysMetAlaLysGlyLysPheAsnValThrLysGlnLysArg 350  
 1001 ATGAGGCGATGCTAAGAAAGAAATTAATGATGACCAAAAGAGCATAGA 1050  
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367

```

1051 AATCAATGATACATGGGGGTCGCCAATAAGCTTACCCCTCAGCAAGTTTC 1100
367 rgllysglythrCysIleGlyAlaProProSerHisGlnHisLeu 384
1101 CGGGAAGGAACTGCATAGGAAAGCTCCCATCCCAACCAACCCCTT 1150
384 ystYserThrValValTyrGlnAlaSerGlnAsnGlnTyrLeuVal 400
1151 GCATAGTACTGGTTTATGAACAGCGCTCAGAAATACATGATTAGTA 1200
401 ProGlyTyrAsnArgTTrpAlaCysAsnThrGlyLeuThrProCysVa 417
1201 CCTGGTTATACAGGTGGTGGCATGCAATACATGGTTAACCCCTGTGT 1250
417 lserThrSerValPheAsnGlnSerLysAspLeuValMetValGlnI 434
1251 TTCACACTAGTCTTTAACCAATCCAAAGATTCTGTCTCATGTCCTCAA 1300
434 lvalProArgValTyrTyrHisProGlnGluValValLeuAspGluTyr 450
1301 TCGTCCCCGAGTGTACTACCATCTCTGAGAAAGTGTCTGTATGATAT 1350
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
1351 GACTATCGGTAAACCGACCGAAAGAGAACCCGTATCCCTTACCCTAGC 1400
467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 484
1401 TGTATGCTCGGATTAGGAGCGCCCTTGGGTAGGAACAGGACACTG 1450
484 lAlaLeuIleThrGlyProGlnGlnLeuGlyGlyLeuGlyLeuHis 500
1451 CCTGATACACAGACACAGCAGCTAGAGAAAGACTTGGTGAAGCTATCAT 1500
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerValSerAsnLe 517
1501 GCGGCCATGACAGAGATCTCCGAGCCTTAGAGAGTCTGTAGCAACCT 1550
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
1551 AGAAGACTCCCTGACTCTTGTCTGAGAGTGGTTCTACCAACCGGAGG 1600
534 lYleuAspLeuLeuPheLeuArgGluGlyLeuGlyCysAlaAlaLeuLys 550
1601 GATTAGATCTGCTGTTCTTAAGAGAGGTGGTATGTGTGAGCCTTAAA 1650
551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerNe 567
1651 GAAGAAATGTTGCTTCTATGTATGATCACTCAGAGCCATCAGACATCCAT 1700
567 lAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 584
1701 GACCAAGCTTAGAAGAAAGTTAGAGAGCGCTCAAGAGGAAAGAGAGCTG 1750
584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600
1751 ACCAGGGGGTGTGAAGGATGGTTCACACAGGCTCTGATGACACAC 1800
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 617
1801 CTGCTTTCGCTCTGACGGGACCGCTAGTAGTCTGCTCTGTACTTAC 1850
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 634
1851 AGTTGGGCTTGCTTAATATAGGTTTGTCTCTTGTAGAGAATGAG 1900
634 AlSerAlaValGlnIleMetValLeuArgGln 645
1901 TAAAGTCACTCCAGATCATGTACTTAGACACAGC 1935
seq_name: gb_v1:PEN288590

```

```

seq_documentation_block:
LOCUS      PEN288590      1996 bp      RNA      VRL      29-AUG-2000
DEFINITION Porcine endogenous retrovirus PERV-B4 env gene for envelope
            protein, genomic RNA.
ACCESSION  AJ288590
VERSION    AJ288590.1 GI:7414449
KEYWORDS   env gene; envelope protein.
SOURCE     porcine endogenous retrovirus.
ORGANISM   Porcine endogenous retrovirus
            Viruses; Retroviral viruses; Retroviridae; Mammalian type C
            retroviruses; 1-Mammalian type C virus group.
REFERENCE  1 (bases 1 to 1996)
AUTHORS    Bosch, S., Arnold, C. and Jestin, A.
TITLE      Study of full-length porcine endogenous retrovirus genomes with
            envelope gene polymorphism in a specific-pathogen-free large white
            swine herd
JOURNAL    J. Virol. 74 (18), 8575-8581 (2000)
MEDLINE    20411432
REFERENCE  2 (bases 1 to 1996)
AUTHORS    Bosch, S.
TITLE      Direct Submission
JOURNAL    Submitted (28-MAR-2000) Bosch S., Molecular Biology Department,
            AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE
FEATURES   location/qualifiers
            1..1996
            /organism="porcine endogenous retrovirus"
            /specific_host="Sus scrofa"
            /db_xref="taxon:51673"
            /clone="PERV-B4"
            /note="French Large White SPF-pig"
            1..1215
            /gene="env"
            1..1215
            /gene="env"
            /note="truncated, not functional"
            /codon_start=1
            /product="envelope protein"
            /protein_id="CAB86209.1"
            /db_xref="GI:7414450"
            /translation="MRPSSRRLLPTRGGEPRRLRIPSPASTAMFLTLTVPQASSK
            RLIDSSDHPRLSLITWLLIDPDGVYVNSRGAAPRGTMPELHFLRLINPAVKSTP
            PNIVRSYGFCCGTEKEERECGGRSFCGRMSCVTSNODGMKMPISLIDPAVKSFVN
            SGPEKIVMLELYDKSCSPSDLDYKRISTFEKROQNIQKIMGMSMGIVFYKGGGA
            GSTLTLLRIETHGTPVAVGPNIVLAEDPRLPEPHNLVPLTSLRPDITQPDN
            STTGLPTNTPRNSPGVPVYTGQRLFSLLOGAFOALNSTDPASSSCMLTSSGPPY
            EGMAREGKENVTRKHRDQCWMSRNKLTILEVSGKFTCI GKAPRSHQHLNSTMYED
            ALENOYLVPQYN"
BASE COUNT 521 a 499 c 503 g 472 t 1 others
ORIGIN
align_scores:
  Quality: 3130.50      Length: 646
  Ratio: 5.066          Gaps: 2
  Percent Similarity: 95.666      Percent Identity: 91.176
alignment_block:
US-09-171-553b-10 x PEN288590
Align seg 1/1 to: PEN288590 from: 1 to: 1996
1 MethSProthrlLeuSerTTrpArgHisLeuProThrArgGlyGluPr 17
1 ATGGCTCCACAGTCACAGCCGCGCCACTCCGACTCGGGGTGAGAGCC 50
17 OlyArgLeuArgLLeProLeuSerPheAlaSerIleAlaTrpPheLeu 34
51 GAAGAAGCTAGAAATCCCTTAAGCTTCCCTTCATCGCTGGTTCCTTA 100
34 hrlLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSer 50
101 CTCTAACAGTGAAGTCCCGAGCGCTAGTAGTAACGCTTATAGACAGCT 150

```

```

51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspThr 67
151 GACCCCATAGACCTTATCCCTTATGCGCTGATTTATGACCCCTGATAC 200
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
201 GGGGTGTCTGTATATAGCTGAGGTGTGCTCTAGAGGCACTGGT 250
84 rPProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
251 GGCCTGAACCTGATTTCTGCTCCGATTTGATTAAACCCGCTTTAAAAAC 300
101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyThr 117
301 ACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTATTGCTGCCAAGCAC 350
117 rGluLysGluLysTrpCysGlyGlySerGlyGluSerPheCysArgArgT 134
351 AGAGAAAGAGAAATCTGTGGGGTTCCAGGGAACTCTTCTGTGGAGAT 400
134 rPSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150
401 GGAGCTGGGTACCTCCACAGATGAGACGTGAAATGGCCGATCTCTCTC 450
151 GluAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrp 167
451 CAGAGCCGGTAAATTCCTCTTGTCAATCCGGCCCGGCAAGTCAAA 500
167 SmeKelLysLeuTrpLysAspLysSerCysSerProSerAspLeuAspT 184
501 AGTGTGGAATATATAAGATAAGAGCTGCTCCCATCAAGACTTAAAT 550
184 rYLeuLysIleSerPheThrGluLysGlyLysGluAsnIleGluLys 200
551 ATNTAAAGATAGTTTCTACTGAAAAAGAAACAGGGAATATTCAAAAG 600
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGlyGly 217
601 TGGATTAATGTTATGAGCTGGGAATAGTTTTTATTAATATGAGGGGG 650
217 rYAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 234
651 AGCAGGGGTCACCTTAATCTATCCCTTAAAGATAGAGACGGGCAAGAA 700
234 rProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 250
701 CCCCTGTGGAGTGGGAGCCACACATAGTACTGGCTGAACAGAGACCCCG 750
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeu 267
751 GCCCTGGAGCCACACAACTTCCCGGTGCCCAAGTTAACTTCGCGCG 800
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 284
801 GCCCTGACATTAACACAGCCGCTGCAACAGTACACCTGATTCCTTA 850
284 hrAsnThrProArgAsnSerProGlyValProValLysThrGluGln 300
851 CCAACAGCCCTTAAGAACTCCCGGGTGTCTCTGTTAAGACGGACAGAGA 900
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317
901 CTCTTCAGTCTCATCCAGGAGCTTCCAAAGCCATCACTCCACGAGACC 950
317 cAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 334
951 CGATGCCAGTCTCTCTGTTGGCTTGTCTATCCTCAGGGGCTCTCTATT 1000
334 rYrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350
1001 ATGAGGGAATGGCTAGAGAAAGGAAATTCATGTGACCAAGACCATAGA 1050
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367

```

---

```

1051 GATCAATGTATCATGGGGGTCCCGAAATTAAGCTCACCTCATGAACTTTC 1100
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 384
1101 CGGGAAGGGGACATGATAGGAAAGCTCCCATCCACCAACACTTTT 1150
384 yTrpSerThrValAlaTrpGluGlnAlaSerGluAsnGlnTrpLeuVal 400
1151 GCAATAGTACTATGTTTATGAGCAGCCCTTAGAAATACATGATTTAGTA 1200
401 ProGlyTrpAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCys 417
1201 CCGTGTATTAATAGTATGGGATGATGCAATACAGGTTAACTCCCTGT 1250
417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 434
1251 TTCCACCACAGCTTTAACCATAATCCAAAGATTTCTGTATGATGTCAC 1300
434 lValProArgValTrpTrpHisProGluGluValValLeuAspGluTr 450
1301 TCGTCCCTGGGTACACTACATCTGAGAGATAGTCTCTCATGAGTTT 1350
451 AspTrpArgTrpAsnArgProLysArgGluProValSerLeuThrLeu 467
1351 GACTATCGGCTACCCGCTCCAAAGAGACCCGTGACCTTAAACCTAG 1400
467 lValMetLeuGlyLeuGlyThrAlaValGlyAlaGlyThrGlyThrAla 483
1401 CCGTTATTAAGATTAGGATGAGT...GGTGTGGGAGAGGACAGCT 1447
484 AlaLeuIleThrGlyProGlnGlnLeuGlyLysGlyGluGluLeu 500
1448 GCCCTGATACAGAGCAGACAGCAGCTAGAAAGAAAGCTTGGTGAAGT 1497
500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsn 517
1498 TCGGCGCATTAACAGAAATCTCCGAGCCCTAGAGAAATGTTAAACAA 1547
517 eGluGluLysLeuThrSerLeuSerGluValValLeuGlnAsnArg 533
1548 TAGAAGAAATCCCGATTTCTGTGTGAAGTGGTTCTTACAGAACTGGA 1597
534 GlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeu 550
1598 GGATTAATCTGCTGTTCTTAAAGAAAGTGGGTTATGTGCACCTTAAG 1647
550 sGluGluCysCysPheTrpValAspHisSerGlyAlaIleArgAspSer 567
1648 AGAAGATGTTGCTCTATGTGATCAGTACAGAGCATCAGAGACTCA 1697
567 eCAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
1698 TGAAGCAAGCTTACAGAAAGTTTAGAGGAGCTCGAAGGAGAAAGAG 1747
584 AspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThr 600
1748 GACCAGGGGTGTTTAAAGATGTTCAACAGGTTCTCTTGGATGACCA 1797
600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeu 617
1798 CCGTCTTCTGCTCTGATGGAGCTCTAGTGGCTCGCTCGTGTGCTTA 1847
617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
1848 CAGTTGGGCTTGTCTTAATTAATAGGTTGTGCTTGTAGAGAAAGA 1897
634 ValSerAlaValGlnIleMetValLeuArgGlnGln 645
1898 GTGAGTGCAAGTCCAGATCATGTACTTAAACACAA 1933

```

seq\_name: gb\_v1: PEN288591

seq documentation block:

LOCUS PEN288591 2002 bp RNA VRL 29-AUG-2000

DEFINITION Porcine endogenous retrovirus PERV-B6 env gene for envelope protein, genomic RNA.

ACCESSION AJ288591

VERSION AJ288591.1 GI:7414451

KEYWORDS env gene; envelope protein; porcine endogenous retrovirus; porcine endogenous retrovirus viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group.

SOURCE ORGANISM

REFERENCE 1 (bases 1 to 2002)

AUTHORS Bosch, S., Arnaud, C. and Jestin, A.

TITLE Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific-pathogen-free large white swine herd

JOURNAL J. Virol. 74 (18), 8575-8581 (2000)

MEDLINE 20411432

REFERENCE 2 (bases 1 to 2002)

AUTHORS Bosch, S.

TITLE Direct Submission

JOURNAL Submitted (28-MAR-2000) Bosch S., Molecular Biology Department, AFSSA-Zoopole, Zoopole Les Croix B.P.53, Ploufragan 22440, FRANCE

FEATURES

source

1..2002

location/Qualifiers

/organism="porcine endogenous retrovirus"

/specific\_host="Sus scrofa"

/db\_xref="taxon:61673"

/clone="PERV-B6"

/note="French Large White SPF-pig"

1..864

/gene="env"

1..864

/gene="env"

/note="truncated, not functional"

/codon\_start=1

/product="envelope protein"

/protein\_id="CAB86210.1"

/db\_xref="GI:7414452"

/translation="MGPRSSRRLLPMWGEPRKRLPLSPASIAMFLTPNSPOASSK RLIDSSDPHRLSLTWLIDPDGVTVNSIRGVAPRGWMPLELRLINPAVKSTP PNLVRSYGFECGCGEKESGSGRSFCGRMCVTSNMGDMPISLDRYKFSFVN SGPGKYVMELYKDKGSPSLDLYLTKISFTEKKOBNKIMKNGMSGIVFYKYGGA GSRLIRLRERGTETPPVAVGNIVIAEGBPALPEPHNDPPOLTLRPDITTPPDN STTGLISTYOH"

BASE COUNT 531 a 495 c 503 g 472 t 1 others

ORIGIN

alignment\_scores:

Quality: 3056.50 Length: 646

Ratio: 4.978 Gaps: 2

Percent Similarity: 95.046 Percent Identity: 90.402

alignment\_block:

us-09-171-553b-10 x PEN288591 ..

Align seg 1/1 to: PEN288591 from: 1 to: 2002

1 MetHisProthrLeuSerTrpArgHisLeuProThrArgGlyGluPr 17

17 OlyArgLeuArgGlyLeuProLeuSerPheAlaSerIleAlaTrpPheLeu 34

51 GAAAGATTGAGAACCCCTTAAGCTTCGCTCATCGCTGTTCTCTTA 100

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50

101 CTTTAAAGAGTATCCAGGCCAGTAGTAACGCCCTTATAGACAGTTGC 150

51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67

151 GACCCCATAGACCCCTTATCCCTTACGTGGCTGATTTATGACCCCTGATAC 200

67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84

201 GGGTGTACTGTATTATAGACATCGAGGTGTCCTCCATAGAGCCACTGGT 250

84 rPpGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100

251 GGCCGTGAACCTGATTCCTGCTCCGATGATTAACCCCGCTGTTAAAGC 300

101 ThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117

301 ACACCTCCCAACCTAGTCCGAGTTATGGTTCTATTGCTGCCAGGCAC 350

117 rGluLysGluLysTrpCysGlyGlySerGlyGluSerPheCysArgArgT 134

351 AGAAGAAAGAGAAATACGTGTGGGGTTCAGAGAACTCTTGTGGAGAT 400

134 rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150

401 GGAGCTCGCTACCTCCACAGATGAGACTGGAATATGCGCATCTCTC 450

151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrpL 167

451 CAGGACGGGTAAATTCCTCTTGTCAATTCGGCCCGGCAAGTACAA 500

167 SmetLysLeuLysTrpLysAspLysSerCysSerProSerAspLeuAsp 184

501 AGTATGGAACCTATATTAAGATTAAGACTGCTCCCATCAGACTTGAT 550

184 rLeuLysLysSerPheThrGluLysGlyLysGlnLysIleGlnLys 200

551 ATCTMAAGATTAAGTTCTCCTGAAGAAAGCAAGCAAAATATTTCAAA 600

201 TrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGlyGly 217

601 TGGATTAATGTATGACTGGGCAATAGTTTATTATATATGAGAGGG 650

217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 224

651 AGCAGGCTCCACTTAACCTATGCGCTTAGATAGAGAGCGGAGCAAC 700

234 rProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250

701 CCCCTGTGGCAGTGGGACCCACATAGTACTGCTCAACAGGGGCCCCG 750

251 AlaLeuGluProProHisAsnLeuProValProGluLeuThrSerLeu 267

751 GCCCTGAGCCACACACACACTTGGCGGTGCCCAAGTTAAGCTCGTGG 800

267 gProAspLysThrGlnProProSerAsnSerThrThrGlyLeuLeu 283

801 GCCTGACATTAACACAGCGCGCTGCAACAGTACACCTGAGTTGATTC 850

284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyLys 300

851 ACCAACACGCGCTAGAACTCCCGGTGTTCTGTTAAGACAGACAGAG 900

300 GLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317

901 ACCTTCAGCCTCATCCAGGAGACTTCCAGACATCACTCACCGGAC 950

317 rAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTr 333

951 CCATGCGCACTCTCTTATTTGCTTATGCTTATACCTCAGAGGCTCTTAT 1000

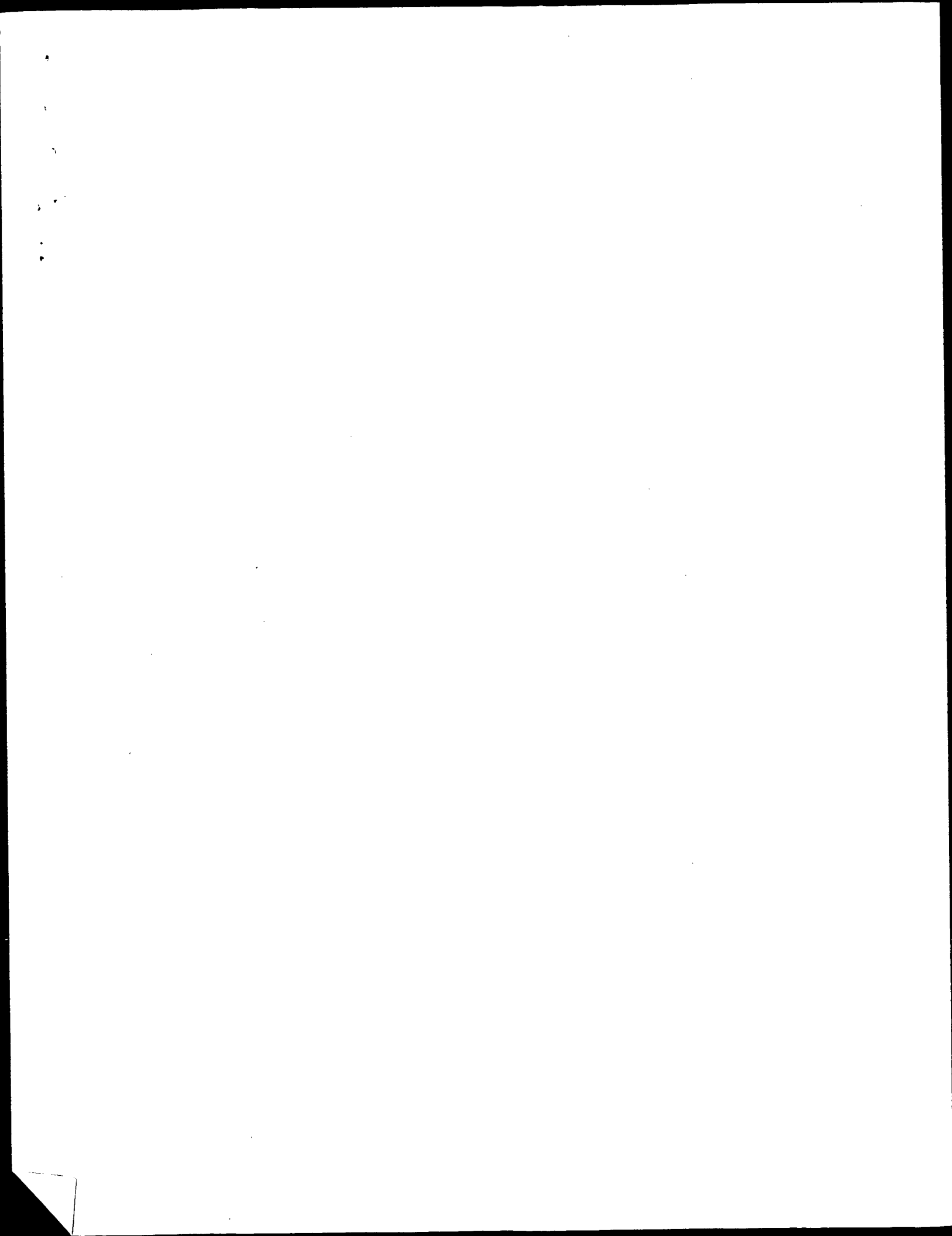
334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 350

1001 TATGAGGAATGGCTAGAGAGGAAATTCAAATGTACCAAGAGAGCAT 1050

350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVal 367

1051 AATCAATGATGATGGGGTCCGAAATTAAGCTCACCTCAGCGAAGTT 1100

```
367  erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383
|||||
1101  CCGGGAAGGGGACATGCATAGAAAGGTCGCCCATCCACACACACCTT 1150
|||||
384  CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400
|||||
1151  TGCATATGTAAGTTATGATGACAGCCCTTAGAAATCATGATTTAGT 1200
|||||
400  LProGlyTyrAsnArgTrrPrrAlaCysAsnThrGlyLeuThrProCysV 417
|||||
1201  ACCGTGTTATACGTGATGGTGGCATCATCTGGGTTAACTCCCTGTG 1250
|||||
417  aLSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433
|||||
1251  TTTCTACACACAGTCTTTAACCAATCCAAAGATTTCGTG.GTCATGTCTCAG 1299
|||||
434  IleValProArgValTyrTyrHisProGluGluValValLeuAspGlnTyr 450
|||||
1300  CTGCTCCCTGGGTACACTACCTACCTGAGAGAAANTCATCGATGAGTA 1349
|||||
450  rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuVal 467
|||||
1350  GGACTATCGGCTAACCCGCTCAAAAAGAGAGCCCGTGACCTTAACCTTAG 1399
|||||
467  IValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483
|||||
1400  CCGTTATACTAGATTAGGATGGCT...GGTGTGGGACAGGACAGAGCT 1446
|||||
484  AlaLeuIleThrGlyProGlnGlnLeuGlyLeuGlyLeuGlyLeuGly 500
|||||
1447  GCCCTGATCACAGACACAGACAGATTAGAAAAAGACTTGGTGAGCTACA 1496
|||||
500  sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517
|||||
1497  TCGCGCATTAACAGAAAGATCTCCGAGCCTTAGAGAAATCTGTTAACACAC 1546
|||||
517  euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
1547  TAGAAGAAATCCCTGACTCT.TTGTCTGAAGTGGTCTTACAGAACTGAGG 1595
|||||
534  GlyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 550
|||||
1596  GGATTAGATCTGCTGTTCTCTAAAAGAGGTGGGTTATGTGACGCCCTTAG 1645
|||||
550  sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567
|||||
1646  AGAAGAATGTGCTTCTATGTGATCACTCAGAGAGCATCAGAGACTCCA 1695
|||||
567. etAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583
|||||
1696  TGGACAAGCTTAGAGAAAGTTAGAGACGTCGCAAGGGAAGAGAGAGCT 1745
|||||
584  AspGlnGlyTrrPrrPheGlnGlyTrrPrrPheAsnArgSerProTrrMetThrTh 600
|||||
1746  GACCAAGGGGTGTTTGAAGAGTGTTCACACAGGTCCTCTGGATGACCA 1795
|||||
600  rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuL 617
|||||
1796  CCTGCTCTCTGCTGATGGAGCCCTAGTGTCTGCTCTGCTGCTGTTA 1845
|||||
617  hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633
|||||
1846  CAGTTGGGCTTGTAAATTAATAGTTATTTGCTTGTAGAGAAATGA 1895
|||||
634  ValSerAlaValAlaGlnIleMetValLeuArgGlnGln 645
|||||
1896  GTGAGTGCAGTCCAGATCATGTACTTAGACACAGCC 1931
|||||
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:56:21 ; Search time 10411.9 Seconds  
(without alignments)  
28.520 Million cell updates/sec

Title: US-09-171-553B-14

Perfect score: 18

Sequence: 1 ccacagtcgtacaccacy 18

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_om.\*

20: em\_or.\*

21: em\_ov.\*

22: em\_pat.\*

23: em\_ph.\*

24: em\_pl.\*

25: em\_ro.\*

26: em\_sts.\*

27: em\_sy.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htgo\_hum.\*

31: em\_htgo\_inv.\*

32: em\_htgo\_rod.\*

33: em\_htgo\_hum.\*

34: em\_htg\_inv.\*

35: em\_htg\_rod.\*

36: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

8

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	100.0	590	14	PEN298074
C 2	18	100.0	633	6	AJ298074 Porcine e
C 3	18	100.0	704	6	AX052622 Sequence
C 4	18	100.0	746	14	AX052621 Sequence
C 5	18	100.0	1042	14	AJ298073 Porcine e
C 6	18	100.0	3025	4	Y17012 Porcine end
C 7	18	100.0	3482	6	AF147808
C 8	18	100.0	3482	6	AX002804
C 9	18	100.0	4918	14	PERENV2
C 10	18	100.0	6076	6	AX052637 Sequence
C 11	18	100.0	6076	6	AX052636
C 12	18	100.0	7333	4	AX052647 Sequence
C 13	18	100.0	7333	6	AF038601
C 14	18	100.0	7362	6	AR130474
C 15	18	100.0	7873	6	AX052634 Sequence
C 16	18	100.0	8060	6	AX052638
C 17	18	100.0	8132	4	AR130473 Sequence
C 18	18	100.0	8132	4	AF038599
C 19	18	100.0	8132	4	AF038600
C 20	18	100.0	8132	6	AR130475
C 21	18	100.0	8209	6	A66552
C 22	18	100.0	8209	6	A66553
C 23	18	100.0	8750	14	PEN133818
C 24	18	100.0	8763	14	AJ293657 Porcine e
C 25	18	100.0	8849	14	AJ133817
C 26	18	100.0	8918	14	PEN133816
C 27	17	94.4	18	6	AX118774
C 28	17	94.4	20	6	AR130533
C 29	16.4	91.1	595	14	PEN298075
C 30	16.4	91.1	610	4	AF286463
C 31	16.4	91.1	1387	10	AF161301
C 32	16.4	91.1	1491	10	AF106919
C 33	16.4	91.1	1878	3	POV278955
C 34	16.4	91.1	2945	3	AF177661
C 35	16.4	91.1	8918	14	PEN293656
C 36	16.4	91.1	90251	2	AC074230
C 37	16.4	91.1	110000	2	AC073804_4
C 38	16.4	91.1	110000	2	AC087331_1
C 39	16.4	91.1	110000	2	AC087782_1
C 40	16.4	91.1	110000	2	AC090843_2
C 41	16.4	91.1	123835	10	MM303011
C 42	16.4	91.1	148588	2	AC073827
C 43	16.4	91.1	165563	2	AC083752
C 44	16.4	91.1	170906	2	AC084314
C 45	16.4	91.1	174533	2	AC073718

## ALIGNMENTS

RESULT	1	Pen298074/c
LOCUS	Pen298074	Porcine endogenous retrovirus type C proviral LTR sequence (strain 17-JUL-2001)
DEFINITION	PERV-B(43)), 590 bp.	
ACCESSION	AJ298074	
VERSION	AJ298074.1	GI:14970733
KEYWORDS	long terminal repeat; LTR.	
SOURCE	porcine endogenous retrovirus.	
ORGANISM	porcine endogenous retrovirus	
REFERENCE	Viruses; Retrov. 1988; 10:1-10. Mammalian type C virus group.	
AUTHORS	1 (bases 1 to 590)	
TITLE	Czauderna,F., Fischer,N., Bolter,K., Kurth,R. and Tonjes,R.R.	
JOURNAL	Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells	
MEDLINE	J. Virol. 74 (9), 4028-4038 (2000)	
REFERENCE	20219394	
AUTHORS	2 (bases 1 to 590)	
	Scheef,G., Fischer,N., Krach,U. and Tonjes,R.R.	

TITLE The Number of a U3 Repeat Box Acting as an Enhancer in Long Terminal Repeats of Polytropic Replication-Competent Porcine Endogenous Retroviruses Dynamically Fluctuates during Serial Virus Passages in Human Cells

JOURNAL MEDLINE 21329475

REFERENCE 3 (bases 1 to 590)

AUTHORS Toenjes, R.R.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-2000) Toenjes R.R., Genetically Engineered Vaccines, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-59, 63225, GERMANY

FEATURES Location/Qualifiers

source 1..590

/organism="porcine endogenous retrovirus"

/proviral

/strain="PERV-B(43)"

/db\_xref="taxon:61673"

1..590

1..428

/note="U3 region"

400..405

misc\_feature 429

/note="Cap site"

429..507

/note="R region"

508..590

/note="U5 region"

BASE COUNT 146 a 154 c 140 g 150 t

ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 590;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18

|||||

Db 466 CCACAGTCGTACACCAG 449

RESULT 2

AX052622/c

LOCUS

DEFINITION Sequence 633 bp DNA PAT 12-JAN-2001

ACCESSION Sequence 6 from Patent WO0071726.

VERSION AX052622.1 GI:12226812

KEYWORDS

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

Viruses; Retrovirus; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 633)

Federspiel, M.J.

Methods to inhibit infectious agent transmission during

xenotransplantation

Patent: WO 0071726-A 7 30-NOV-2000;

Mayo Medical Ventures (US)

FEATURES Location/Qualifiers

source 1..633

/organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"

BASE COUNT 170 a 125 c 147 g 191 t

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 633;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18

|||||

Db 497 CCACAGTCGTACACCAG 480

RESULT 3

AX052621/c

LOCUS

DEFINITION Sequence 6 from Patent WO0071726.

ACCESSION AX052621

VERSION AX052621.1 GI:12226811

KEYWORDS

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

Viruses; Retrovirus; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 704)

Federspiel, M.J.

Methods to inhibit infectious agent transmission during

xenotransplantation

Patent: WO 0071726-A 6 30-NOV-2000;

Mayo Medical Ventures (US)

FEATURES Location/Qualifiers

source 1..704

/organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"

BASE COUNT 187 a 162 c 158 g 197 t

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 704;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18

|||||

Db 568 CCACAGTCGTACACCAG 551

RESULT 4

PEN298073/c

LOCUS

DEFINITION Porcine endogenous retrovirus type C proviral LTR sequence (strain

PERV-B(43)), 746 bp.

ACCESSION AJ298073

VERSION AJ298073.1 GI:14970732

KEYWORDS

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

Viruses; Retrovirus; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 746)

Czaderna, F., Fischer, N., Boller, K., Kurth, R. and Tonjes, R.R.

Establishment and characterization of molecular clones of porcine

endogenous retroviruses replicating on human cells

J. Virol. 74 (9), 4028-4038 (2000)

20219394

REFERENCE 2 (bases 1 to 746)

Scheef, G., Fischer, N., Krach, U. and Tonjes, R.R.

The Number of a U3 Repeat Box Acting as an Enhancer in Long

Terminal Repeats of Polytropic Replication-Competent Porcine

Endogenous Retroviruses Dynamically Fluctuates during Serial Virus

Passages in Human Cells

J. Virol. 75 (15), 6933-6940 (2001)

21329475

REFERENCE 3 (bases 1 to 746)

Toenjes, R.R.

Direct Submission

TITLE Submitted (05-DEC-2000) Toenjes R.R., Genetically Engineered

Vaccines, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-59, 63225,

GERMANY

FEATURES Location/Qualifiers

source 1..746

/organism="porcine endogenous retrovirus"

/proviral

/strain="PERV-B(43)"



```

/db_xref="taxon:61673"
LTR 1..746
misc_feature 1..584
TATA_signal /note="U3 region"
misc_feature 556..561
misc_feature 585
misc_feature /note="Cap site"
misc_feature 585..663
misc_feature /note="R region"
misc_feature 664..746
misc_feature /note="U5 region"
BASE COUNT 182 a 186 c 184 g 194 t
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 746;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 622 CCACAGTCGTACACCACG 605

RESULT 5
PERY17012/c
LOCUS PERY17012 1042 bp DNA VRL 14-APR-2000
DEFINITION Porcine endogenous retrovirus DNA for long terminal repeat.
ACCESSION Y17012
VERSION Y17012.1 GI:3005984
KEYWORDS long terminal repeat; LTR.
SOURCE porcine endogenous retrovirus.
ORGANISM Porcine endogenous retrovirus.
REFERENCE 1 (bases 1 to 1042)
AUTHORS Czauderna,F.; Fischer,N., Boller,K., Kurth,R. and Tonjes,R.R.
TITLE Establishment and characterization of molecular clones of porcine
endogenous retroviruses replicating on human cells
J. Virol. 74 (9), 4028-4038 (2000)
20219394
REFERENCE 2 (bases 1 to 1042)
AUTHORS Tonjes,R.R.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-1998) R.R. Tonjes, Paul-Ehrlich Institut,
Paul-Ehrlich Str. 51-59, P.O.Box 1740, D-63207 Langen, FRG
LOCATION/Qualifiers
1. 1042
FEATURES
source
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
/tissue_type="kidney"
/cell_line="PK15"
1..1042
LTR 214 a 270 c 289 g 269 t
BASE COUNT
ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 1042;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 503 CCACAGTCGTACACCACG 486

RESULT 6
AF147808/c
LOCUS AF147808 3025 bp DNA MAM 27-APR-2000
DEFINITION Sus scrofa domestica porcine endogenous retrovirus type C gag-pol
precursor, gene, partial cds.
ACCESSION AF147808
VERSION AF147808.1 GI:7650139

```

```

KEYWORDS domestic pig.
SOURCE Sus scrofa domestica
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 3025)
AUTHORS Blusch,J.H., Seelmeir,S. and von der Helm,K.
TITLE Evidence for extreme sequence homogeneity of the PERV A/B 5'
LTR/leader/gag/prot region by predicted targeted amplification from
a Du x DL pig
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3025)
AUTHORS Blusch,J.H.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-1999) Virology, Max-von-Pettenkofer-Institute,
Pettenkofer-Str. 9a, Munich 80336, Germany
FEATURES
source
Location/Qualifiers
1..3025
/organism="Sus scrofa domestica"
/strain="breed Duroc x German Landbreed, Du x DL3791"
/sub_species="domestica"
/db_xref="taxon:9825"
1..3025
repeat_region
/note="PERV A/B; porcine endogenous retrovirus type C"
/rpt_family="perv"
/rpt_type="dispersed"
587..605
primer_bind
1032..>3025
CDS
/note="gag protein produced by proteolytic cleavage of the
mature protease out of the precursor"
/codon_start=1
/transl_except="(pos:2604..2606,aa:Gln)"
/product="gag-pol precursor"
/protein_id="AAF65926.1"
/db_xref="GI:7650140"
/translation="MGQTVTTPLSLTLDHWTEVRSRAHNLVSQVKKGPWOTFCASEWP
TFDVGWPSQGTNSEITILAVKAIIFOTGSPSHDQEPYILTWQDLAEDPPVKKWLN
KPRKPGRIILALGEKNKSAEKVSPRIYPLEEPTWPEQVPVPPYPAGGAVRG
PSAPGAPVVEGPAAGTRSRRCATPERTDEAILPLRTGTPMPGQQLQLOIWFPS
ADLYNWKTHPPFSDPQRLTGLVLSHSDTDDCCQLLOTLTTEERERILLEA
RKNVPGADRTQLQNEIDMGFPLTRPDYNTAEGRESKIYRQALVAGLRGARRP
TNLAQVRYMQGPNPSPVFLERLMEAFRRFPDPTSEAKASVALAPIGOSALDIR
KKLQRLGLEQAEALDLVREAKVYVRETEKEQKEREEREERDRQENLT
KILAAVVEGKSSRERDRFKIRSGPROSGNLGNTPLDKOCAYCKEKGHWARNCPK
KNGKGPVLALEDKQGRGSDPLPEPRVTLKVEGQPVFVLDTCAGHSVLLPLGK
LKEKSWMGATGQROYPWTRTRTDVLDGVRVTHSLVIPCVPVLLGLLTKMGAG
ISFEQGRPEVSNNKPTITVLTQLDDEYRLYSPQVKPDQD"
1032..2603
mat_peptide
/product="gag protein"
2604..>3025
mat_peptide
/product="aspartic protease"
BASE COUNT 743 a 749 c 849 g 684 t
ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccacg 18
|||||
Db 461 CCACAGTCGTACACCACG 444

RESULT 7
AX002804/c
LOCUS AX002804 3482 bp DNA PAT 21-AUG-2000
DEFINITION Sequence 3 from Patent WO9853104.
ACCESSION AX002804
VERSION AX002804.1 GI:9885132,
KEYWORDS porcine endogenous retrovirus.
SOURCE

```

ORGANISM porcine endogenous retrovirus  
viruses; Retrovirdae; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 3482)  
AUTHORS Stoye,J.P. and Weiss,R.A.  
TITLE Detection of retroviral subtypes based upon envelope specific sequences

JOURNAL Patent: WO 9853104-A 3 26-NOV-1998;  
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)

FEATURES  
source  
1..3482  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

BASE COUNT 927 a 854 c 867 g 834 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 3482;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
|||||

Db 3439 CCACAGTCGTACACCAG 3422

RESULT 8  
LOCUS PERENV2 3482 bp RNA VRL 14-OCT-1997  
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.  
ACCESSION Y12239  
VERSION Y12239.1 GI:2576324  
KEYWORDS env gene.  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
Viruses; Retrovirdae; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 3482)  
AUTHORS Le Tissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.  
TITLE Two sets of human-tropic pig retrovirus  
JOURNAL Nature 389 (6652), 681-682 (1997)  
MEDLINE 97478526  
REFERENCE 2 (bases 1 to 3482)  
AUTHORS Stoye,J.P.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical  
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK  
REMARK revised by author 16-SEP-97

FEATURES  
source  
1..3482  
/organism="porcine endogenous retrovirus"  
/specific\_host="Sus scrofa"  
/db\_xref="taxon:61673"  
/note="PERV-B"  
911..2884  
/gene="env"  
911..2884  
/gene="env"  
/codon\_start=1  
/protein\_id="CAA72928.1"  
/db\_xref="GI:2576325"  
/db\_xref="SPTREMBL:O41173"  
/translation="MHPTLSWRHLPTRGGEKRLRPLSFASIAFWLTITTPQASSK  
RLIDSNPHRLSLTLWLIIDPTGVNSTRGVAPRGTWMPHLFCLRLINPAVKSTP  
PNLVRSGFYCCPTKEKYCGGSGFCRRSCVTSDGDMKWPISLODRVYKFSVP  
SGPKYKVMKLYDKSCSPDLDYLKISTEKGQENIQKWNMGSLVYKYGGGA  
GSILTLRLTETGTPVAVGPKVLAEGPPALPEPHNLVPQLTSLRPDIQPPSN  
GTGLIPTNTPRSPGVVTKGRLFLSQGAQAINSDPDATSCWLCLSGGPPY  
EGMAKEGFNVTKEHRNQCWTGSKNLTLEVSQKTCIGKAPPSHQHLCYSTVVEQ  
ASNQYLVPYGNRWACNTGLTPTCVSTVFNSQKDFCMVQIVPRVYVHPEAVLDEY  
DYRNRKPRPSVLTIAVLMGLQTAVGVTGTAALITGPQLEKGLHGAAMTDLR  
ALRESVNLSESTLSSEVLQNRGLDILLFLREGGLCAALKKECCFYVDHSGAIRD  
MSKLRERLRRRRERADQWFEQFNFSFMTLLLSALTGPLVLLILLTVGPCLIN

BASE COUNT 927 a 854 c 867 g 834 t  
ORIGIN

Query Match 100.0%; Score 18; DB 14; Length 3482;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
|||||

Db 3439 CCACAGTCGTACACCAG 3422

RESULT 9  
LOCUS AX052637/c 4918 bp DNA PAT 12-JAN-2001  
DEFINITION Sequence 22 from Patent WO0071726.  
ACCESSION AX052637  
VERSION AX052637.1 GI:12226827  
KEYWORDS  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
Viruses; Retrovirdae; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 4918)  
AUTHORS Federspiel,M.J.  
TITLE Methods to inhibit infectious agent transmission during  
xenotransplantation  
JOURNAL Patent: WO 0071726-A 22 30-NOV-2000;  
Mayo Medical Ventures (US)  
FEATURES  
source  
1..4918  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

BASE COUNT 1382 a 1173 c 1183 g 1180 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 4918;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
|||||

Db 4781 CCACAGTCGTACACCAG 4764

RESULT 10  
LOCUS AX052636/c 6076 bp DNA PAT 12-JAN-2001  
DEFINITION Sequence 21 from Patent WO0071726.  
ACCESSION AX052636  
VERSION AX052636.1 GI:12226826  
KEYWORDS  
SOURCE porcine endogenous retrovirus.  
ORGANISM porcine endogenous retrovirus  
Viruses; Retrovirdae; Retroviridae; Mammalian type C  
retroviruses; 1-Mammalian type C virus group.

REFERENCE 1 (bases 1 to 6076)  
AUTHORS Federspiel,M.J.  
TITLE Methods to inhibit infectious agent transmission during  
xenotransplantation  
JOURNAL Patent: WO 0071726-A 21 30-NOV-2000;  
Mayo Medical Ventures (US)  
FEATURES  
source  
1..6076  
/organism="porcine endogenous retrovirus"  
/db\_xref="taxon:61673"

BASE COUNT 1613 a 1512 c 1609 g 1342 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 6076;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
 |||||  
 Db 617 CCACAGTCGTACACCAG 600

RESULT 11  
 AX052647/c  
 LOCUS  
 DEFINITION Sequence 32 from Patent WO0071726.  
 ACCESSION AX052647  
 VERSION AX052647.1 GI:12226837  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM  
 Viruses; Retroviral viruses; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 REFERENCE 1 (bases 1 to 6076)  
 AUTHORS Federspiel, M.J.  
 TITLE Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 JOURNAL Patent: WO 0071726-A 32 30-NOV-2000;  
 Mayo Medical Ventures (US)  
 FEATURES Location/Qualifiers  
 source  
 1. 6076  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"  
 BASE COUNT 1619 a 1505 c 1598 g 1354 t  
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 6076;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
 |||||  
 Db 617 CCACAGTCGTACACCAG 600

RESULT 12  
 AF038601/c  
 LOCUS  
 DEFINITION Sus scrofa porcine endogenous retrovirus ERV-PK15 mRNA, complete  
 sequence.  
 ACCESSION AF038601  
 VERSION AF038601.1 GI:3133304  
 KEYWORDS  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 7333)  
 AUTHORS Akiyoshi, D.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P. and  
 Fishman, J.A.  
 TITLE Identification of a full-length cDNA for an endogenous retrovirus  
 of miniature swine  
 JOURNAL J. Virol. 72 (5), 4503-4507 (1998)  
 MEDLINE 98216827  
 REFERENCE 2 (bases 1 to 7333)  
 AUTHORS Fishman, J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts  
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REFERENCE 3 (bases 1 to 7333)  
 AUTHORS Fishman, J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts  
 General Hospital, 55 Fruit Street, Boston, MA 02114, USA  
 REMARK Sequence update by submitter

COMMENT On May 16, 1998 this sequence version replaced gi:3116445.  
 FEATURES  
 source  
 1. 7333  
 /organism="Sus scrofa"  
 /strain="miniature swine"  
 /db\_xref="taxon:9823"  
 /cell\_line="PK-15"  
 /note="porcine endogenous retrovirus ERV-PK15; infective  
 virus but represents a naturally occurring transcript"  
 1. 578  
 598. 2172  
 /codon\_start=1  
 /product="gag protein"  
 /protein\_id="AAC16766.1"  
 /db\_xref="GI:3116446"  
 /translation="MGQTVTTPLSLTDHWTVEKRAHNSVQVKGQWTFCASEWP  
 TFDVGPSEGTNSEIILAVKAIIFQGTGSPHQEPYILTWDLAEDPPWPKPWLN  
 KPRKPRILALGKKNHSAEKVPSRIYPIEIEPTWPEQVPVPPVPAQCAVRC  
 PSAPGAPVVGSPAGTSRGATPERTDEIAILLPTTYGPPMGGQLQLOIYWPFS  
 ADLYNWKTHHPFSDPQRLGLVESLMFSHQPTWDCQQLQTLFTTEERILLLEA  
 KKNVPGADGRPTQONEIDMGFPPLTPGWDYNTAEGRESLKIYQALVAGLGRARRP  
 TNLAKREVMPQNPESVFLERLMEAFREFPTPDPTSEAKASVALAFIQQSALDIR  
 KKLQRLGLEQAEALDLVREAEKVYVRETEKEKEKEKEKEKEKEKEKEKEKILIT  
 KILAAVVGKSKSRERDERFKIRSGPROSGNLGNRTPLDKQCAVCKEKGHWARNCKPK  
 KGNKGPVLALEEDKD"  
 2320. 3552  
 /note="truncated protein"  
 /codon\_start=1  
 /product="pol protein"  
 /protein\_id="AAC16767.1"  
 /db\_xref="GI:3133305"  
 /translation="MGATGQRQYPTWTRTRTVLDLGVGRVTHSVFLVIPCVPVLLGRDLL  
 TKGAQISFEGREPEVSNNAITVLTQLDDEYRLSPQVKPDQDQISWLEQFPQAM  
 AETAGMLAKQVPPQVQLQKASATPSVRSQPLSREAREGIWPHVRLIQGLILVPLVQ  
 SPWNTPLLPKPGTNDYRPVQDLREVNKRQYDIHTVTPNPVNLALPERALNVTYL  
 DLKDAFFCLRLHPTSQLPTFEWRDPTGRTGQLTWTLPQGFKNSTPTFDEALHRLD  
 ANFRHQPVQVTLLOYVDDLLAGATKDCLEGTKALLLESLDLGYRASAKKAQICRE  
 VTYLGYSLRGQRWLTEARKTVVQIIPAPTTAKOVREFLGTAGFCRLWIFGFATLAAP  
 LYPLTKKGCGCPQQGK"  
 4803. 6725  
 /note="env protein; N terminus deleted"  
 /codon\_start=1  
 /pseudo  
 6726. 7333  
 1984 a 1788 c 1885 g 1676 t  
 BASE COUNT  
 ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 7333;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccacagtcgtacaccag 18  
 |||||  
 Db 7279 CCACAGTCGTACACCAG 7262

RESULT 13  
 AR130474/c  
 LOCUS  
 DEFINITION Sequence 2 from patent US 6190861.  
 ACCESSION AR130474  
 VERSION AR130474.1 GI:14118799  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 7333)  
 AUTHORS Fishman, J.A.  
 TITLE Molecular sequences of swine retroviruses method of using  
 Patent: US 6190861-A 2 20-FEB-2001;  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1. 7333

BASE COUNT 1984 a 1788 c 1885 g 1676 t  
 ORIGIN /organism="unknown"

Query Match 100.0%; Score 18; DB 6; Length 7333;  
 Best Local Similarity 100.0%; Pred. No. 7.6;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18  
 |||||  
 Db 7279 CCACAGTCGTACACCAG 7262

RESULT 14  
 AX052634/c  
 LOCUS AX052634 7362 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 19 from Patent WO0071726.  
 ACCESSION AX052634  
 VERSION AX052634.1 GI:12226824

KEYWORDS porcine endogenous retrovirus.  
 SOURCE porcine endogenous retrovirus  
 ORGANISM  
 Viruses; Retrovirus; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 7362)  
 AUTHORS Federspiel, M.J.  
 TITLE Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 JOURNAL Patent: WO 0071726-A 19 30-NOV-2000;  
 Mayo Medical Ventures (US)

FEATURES  
 source  
 1..7362  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"

BASE COUNT 1997 a 1821 c 1881 g 1663 t  
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 7362;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18  
 |||||  
 Db 682 CCACAGTCGTACACCAG 665

RESULT 15  
 AX052638/c  
 LOCUS AX052638 7873 bp DNA PAT 12-JAN-2001  
 DEFINITION Sequence 23 from Patent WO0071726.  
 ACCESSION AX052638  
 VERSION AX052638.1 GI:12226828

KEYWORDS porcine endogenous retrovirus.  
 SOURCE porcine endogenous retrovirus  
 ORGANISM  
 Viruses; Retrovirus; Retroviridae; Mammalian type C  
 retroviruses; 1-Mammalian type C virus group.  
 1 (bases 1 to 7873)  
 AUTHORS Federspiel, M.J.  
 TITLE Methods to inhibit infectious agent transmission during  
 xenotransplantation  
 JOURNAL Patent: WO 0071726-A 23 30-NOV-2000;  
 Mayo Medical Ventures (US)

FEATURES  
 source  
 1..7873  
 /organism="porcine endogenous retrovirus"  
 /db\_xref="taxon:61673"

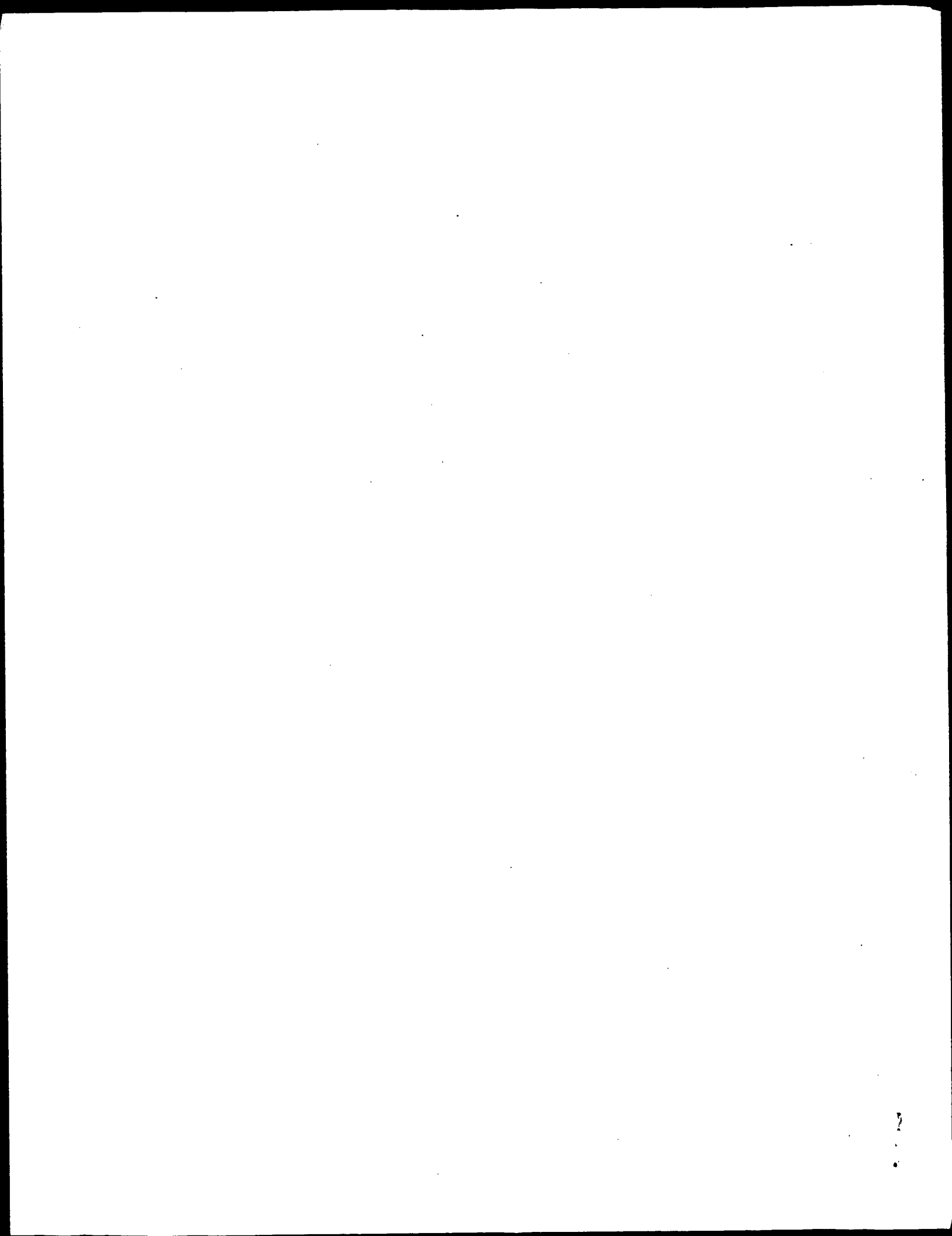
BASE COUNT 2200 a 1914 c 1961 g 1798 t  
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 7873;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccacagtcgtacaccag 18  
 |||||  
 Db 7708 CCACAGTCGTACACCAG 7691

Search completed: February 24, 2002, 02:56:23  
 Job time: 21294 sec







BASE COUNT 922 a 844 c 832 g 722 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3320;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gatggctctctgccccttg 20  
|||||  
Db 2289 GATGGCTCTCCTGCCCTTG 2308

## RESULT 2

AF038601 7333 bp mRNA MAM 21-MAY-1998  
LOCUS Sus scrofa porcine endogenous ERV-PK15 mRNA, complete  
DEFINITION sequence.  
ACCESSION AF038601  
VERSION AF038601.1 GI:3133304  
KEYWORDS  
SOURCE pig.  
ORGANISM Sus scrofa

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 7333)  
AUTHORS Akiyoshi,D.E., Denaro,M., Zhu,H., Greenstein,J.L., Banerjee,P. and Fishman,J.A.

TITLE Identification of a full-length cDNA for an endogenous retrovirus of miniature swine

J. Virol. 72 (5), 4503-4507 (1998)

REFERENCE 98216827

2 (bases 1 to 7333)

AUTHORS Fishman,J.A.

TITLE Direct Submission

Submitted (16-DEC-1997) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA

REFERENCE 3 (bases 1 to 7333)

AUTHORS Fishman,J.A.

TITLE Direct Submission

Submitted (15-MAY-1998) Infectious Disease Unit, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA

SEQUENCE update by submitter

On May 16, 1998 this sequence version replaced gi:3116445.

COMMENT

FEATURES

source

1. 7333

/organism="Sus scrofa"

/strain="miniature swine"

/db\_xref="taxon:9823"

/cell\_line="PK-15"

/note="porcine endogenous retrovirus ERV-PK15; Infective virus but represents a naturally occurring transcript"

1. 578

598..2172

/codon\_start=1

/product="gag protein"

/protein\_id="AAC16766.1"

/db\_xref="GI:3116446"

/translation="MGQTVTPLSLTLDHWTEVRSRAHNSVQVKKGPWQTFCASEWP

TFVGVPSGETFSEIILAVKALIFQPGSHDQEPYILTWDLAEDPPVWKPWLN

KPKPGPRILALGKKNHSAEKVPSRIPYETEEPTWPEPPVPPYPAQGVRG

PSAPGAPVGVGAAGTSPRGATPERTDEIALLPTTYGPPMPGQLOLPYWPSS

ADLYNKNTHNPPESEDPQRLTGLVESIMFSDQTDWDCQQLQTFTTERERILLEA

KKNVPGADGRTQLONEIDMGFFLTPRGWDYNTAEKSLIYRQALVAGLRGARRP

TNLAKRVNMQGPNPPSVFLERLMEAFRRFTPDPTSEAQAVALAFIGSALDIR

KILQLEGLQAEALDLVREAEKVYRRETEBEKEQKEKEEREERDRDRQENLT

KILAAVVEGSSRRERDFKIRSGPRQSGNLGRTPLDKDQACVCKEKGHWARNCPK

KGNGKPKVLAEEDKD"

2320..3552

/note="truncated protein"

/codon\_start=1

/product="pol protein"

/protein\_id="AAC16767.1"

## CDS

/db\_xref="GI:3133305"

/translation="MGATGQROYPWTRTRTVDLGVGRVTHSFLVPIPECVPVLLGRDLL  
TKMQAQSIFEOGRPEVSNKRPITVLTQLDDEYRLYSQVKKPDODIQSWLEQPPQAW  
AETAGMLAKQVPPQVQIOLKASATPSVSRQYPLSREAREGIWPHVRIQOGLVPPVQ  
SPWNTPLLPVRKPTNDYRPVQDLREVNRQVDIHPTVPNPYLLSALPPERNWTVL  
DLKDAFFCLRLHPTSQPLTFEWRDPGTGRGQLTWRLPGQFKNSPTIFDEALHRLD  
ANFRQHQVILLQYVDDLLAGATKQDCLEGTAKALLELSDLYGRASAKKAQICRRE  
VYLYGYSURGGQRWLTEARKKTVVOIPATTTAKQVREFLSTAGFCRWIPGFATLAAP  
LYPLTKRGGCLPQGGK"

## CDS

4803..6725

/note="env protein; N terminus deleted"

/codon\_start=1

/pseudo

6726..7333

BASE COUNT 1984 a 1788 c 1885 g 1676 t

## ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 7333;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gatggctctctgccccttg 20  
|||||  
Db 4442 GATGGCTCTCCTGCCCTTG 4461

## RESULT 3

ARI30474

LOCUS ARI30474 7333 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6190861.

ACCESSION ARI30474

VERSION ARI30474.1 GI:14118799

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 7333)

AUTHORS Fishman,J.A.

TITLE Molecular sequences of swine retroviruses method of using

JOURNAL Patent: US 6190861-A 2 20-FEB-2001;

FEATURES

Location/Qualifiers

Source

1..7333

/organism="unknown"

BASE COUNT 1984 a 1788 c 1885 g 1676 t

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 7333;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gatggctctctgccccttg 20  
|||||  
Db 4442 GATGGCTCTCCTGCCCTTG 4461

## RESULT 4

A66552

LOCUS A66552 8196 bp DNA PAT 29-MAR-1999

DEFINITION Sequence 2 from Patent WO9740167.

ACCESSION A66552

VERSION A66552.1 GI:4538105

KEYWORDS

SOURCE porcine endogenous retrovirus.

ORGANISM porcine endogenous retrovirus

REFERENCE 1 (bases 1 to 8196)

AUTHORS Galbraith,D.N., Haworth,C., Lees,G.M. and Smith,K.T.

TITLE PORCINE RETROVIRUS

JOURNAL Patent: WO 9740167-A 2 30-OCT-1997;

Q ONE BIOTECH LTD (GB)



## FEATURES

source

Location/Qualifiers

1. .8196

/organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"

BASE COUNT 2165 a 2061 c 2146 g 1820 t 4 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 6; Length 8196;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatggctctcgccttg 20

|||||

Db 5226 GATGGCTCTCGCCTTG 5245

RESULT 5

A66553

LOCUS

DEFINITION

A66553

VERSION

A66553.1

KEYWORDS

SOURCE

ORGANISM

porcine endogenous retrovirus.

Viruses; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 8209)

Galbraith, D.N., Haworth, C., Lees, G.M. and Smith, K.T.

PORCINE RETROVIRUS

Patent: WO 9740167-A 3 30-OCT-1997;

Q ONE BIOTECH LTD (GB)

FEATURES

source

Location/Qualifiers

1. .8209

/organism="porcine endogenous retrovirus"

/db\_xref="taxon:61673"

588. .2162

/product="VIRAL CORE PROTEIN"

/protein\_id="CAB39347.1"

/db\_xref="GI:4538107"

/translation="MGQTVPLSLTDHWEVRSRAHNSVQVKKGPWOTFCASEWP

TFDVGWSEGTSEIILAVKAIIFQPGSHDQEPYILTWDLAEDPPWVKPMIN

TKPRPGRIALGPKNKHSAEKVEPSSYLPDRGADLAGTPTCPTPLSSSTGCCG

TSAPPGAPVVEGPAAGTSRSGATPRTDIALPLTYGPPMPGQLOLPWPPSS

ADLNWKNTPPSEDPRLTGIVLSELMFSDQPTDCCQLLOTFTTERERILLLEA

RKNVPGADGPTQLQNEIDMGFPLTRPGDNTAEGRESLIYQALVAGLGAARRP

TNLAKREVWQGNPPSPVSEFLMEAFRRFTPDPTSEAQKASVALAFICQALDIR

KKLQREGLQAEALDLVREAKEYVYRRETEEEKERKEERERERRRROEKNLT

KILAAVVEGKSSRRERDRFRKIRSGPRQSGNLGNRTPLDKDQAYCKEKGHWARNCPK

KGNKGPVLALEEDKQ"

&lt;2163. .5747

/codon\_start=1

/product="REVERSE TRANSCRIPTASE"

/protein\_id="CAB39348.1"

/db\_xref="GI:4538108"

/translation="GRKSDPLPEPRVTLKVEQPVEFLVDTAESVLLQPLGKLKE

KKSWMGATGQRPWTTRVDLGVGRVTHSFLVPECPVLLGRDLTKMAQISF

EQGPVYVNNKPTITVLQDLDEYRLSPQVKPDQDQISWLEQFPQAWAETAGMLA

KQVPQVQLKASATPVSRQYPLSREAREGIWPHVORLIQOGLVPSQVPMWTPLLP

VRKPTNDYRPVODLREVNKVDIHTVPNPNYLLSALPPERNWTVLIDKDAFTCL

RLHPTSLPFAFWDPDGTGRTGQLTWRLPQGFKNSTIFDEALHRDLANFLIHPQ

VTLLQYVDDLLAGATQDCLEGTALLLELSDLYRASAKACICRREVTYLGSLR

GGQWLTARAKTVOIPAPTAKQREFLTGAGFCRLWIPGATLAAPLYPLTKKG

GFSWAPHQAFADA IKKALLSAPALAPDVTKPTVLYVDERKGVARGVLTQTLGWR

PVAYLSKLDLPVAGWPVCLKAATAVAILVKDADKLTLGONITVIAPHALENVRPP

DRWNTNARMTHYQSLLTERTVAPPALNPATLPEETDEPVTTHOCHLIEETGVR

KDLTDIPLTEGLTWTFDCSSVYVEGKRMAGAAVVDGTRTIWASSLPFGTSAGAEIM

ALQALRLAGKSNINITYDSRIAFATAHVGAIYKQGLITSAGREIKNKEETLSLLE

LHLPKRLAITHCPGHQKADLI SRGNQMDRAKAAQAVNLLPIETIPEKPEPRQ

YTLDEMDIKIKIDQFSTPGTCYTSYKIEILPKHGLEVVOQHRLTHLGLTKHLOQL

VRTSPYHVLRLPGVADSVVVKHCVPCQLVNPANSRIIPGKRLRGSHGPAHWEVDTEVK

## CDS

/codon\_start=1

/product="VIRAL ENVELOPE PROTEIN"

/protein\_id="CAB39349.1"

/db\_xref="GI:4538109"

/translation="MHPTLSRRHLPTRGGEKRLRIPLSFASIAFWLTLITTPQASSK

RLIDSNPHRLSLTWLIIDPTGVTNSTRGVAPRGTWMPHELFCRLRLINPAVKSTP

PNLVSRYGFCFTEGTEKERYCGSGESFCRWSCVTSNDGDKWYLSIQDRKVFSEVN

SGPGYKMMKLYKDKSCSPSDLYLKISFTEKTKYKSKVDKWEIGNSFLLYGGAG

STLTIRLETGTETPPVAMPKDLAEQGPALPEPHNLVPQLTSLRDPITPPPSNS

TTCLIPTNTPRNSPGVVKTGQRLSLIOGAFOAINSTDPTDSSCWLCISGPPYE

GNAKERKFNVTKEHNRQCTWGRNKLTLEVSGKGTICGKAPPSHQHLCISYVVEQA

SENQYLVGNRWACNTGLTFCVTSVFNQKDFCVWQIVPRVYIHPVEVVLDEYD

YRYNRPKEPVSILTSLVMGLGTAVGVTGTAALITGPOQLKEGLGELHAAMTEDLRA

LKESVNLSESLTSLSEVVLQNRRLDILLFREGGLCAALKECCFYVDHSGAIRDMS

NKLKKLERRREREADQWGFEGFNRSPWMTLLLSALTGPLVLVLLLLTVGPCLINR

FYAFVPERVSAVQIMVLROQYQGLLSQGETDL"

BASE COUNT 2168 a 2064 c 2154 g 1823 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 8209;

Best Local Similarity 100.0%; Pred. No. 3.7; 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatggctctcgccttg 20

|||||

Db 5240 GATGGCTCTCGCCTTG 5259

RESULT 6

PEN133817

LOCUS

DEFINITION

A133817

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

porcine endogenous retrovirus.

Viruses; Retroviridae; Mammalian type C

retroviruses; 1-Mammalian type C virus group.

1 (bases 1 to 8849)

Czauderna, F., Fischer, N., Boller, K., Krach, U., Kurth, R. and

Toenjes, R.R.

TITLE

Molecular Characterization of Human-tropic and

Replication-competent Porcine Endogenous Retroviruses

Unpublished

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .8849

/organism="porcine endogenous retrovirus"

/strain="type C"

/db\_xref="taxon:61673"

/clone="42"

/note="grown in PK15 (ATCC No. CCL-33) cell line"

1. .668

LTR

misc\_feature

506

/note="cap site"

671. .688

primer\_bind

/note="gly4"

Query Match	100.0%	Score 20;	DB 14;	Length 8849;
Best Local Similarity	100.0%	Pred. No. 3.7;		

```

/translation="GRRGSDPLPERVTLKVEGQVPELVDTGAKSHVLLQPLGKLKD
EKSWNGATGOQYPTWTRTVDLGVGRVTHSFSLVPECPAPLRLDRLTLKQAGISF
EKGKPEVSANRRITVLTILQDDEYLRVSLVFKPDQNIQFWLEQPPQAWAETAGMGLA
KQVPOVIOKASAAAPVROVPLSKAREGRPHVORLQOGLVLPQSWPNTPLLP
VRKPGTNDYRPQDILREVNRVQDILHPTVPNPNLLCALPQORSWYTVLDLKDAPFCL
RLHPTSQPLFAPEWRDPGARGQLTWRLPQGFNSPTIDFADLRDLNFRIOHQ
VILQIVDDLLAGATKQDCLEGTAKLLELSDGLYRASAKAICREVTYLYGYSKR
GTLQRLRAIIRHCPGHQAKDPISRGNMADRVAKQAAGVNLPMIETPKAPEFRQ
YTLDEQELIKIDQFSETPGCTYTDGKEILPHKGELEYVQOIRHLTLGTHLQLOL
VRTSPHYLRPCVADSVVHCPCOLVNPANRIPPCKGLRGLSGHGAHWEVDTEVK
PKYGNKVLVLPVDTFSGWVEAYPTKETSTVWAKKILEEIPFREGIPKVGSDNGPA
FYAQSQGLAKILGIDKWLHCAYRQSSQGVVERMRTIKETIKLTTETGINDWALL
FVLPFRVNTPOGLTYELLYGGPPLAEIAFAHSADVLLSQPLFRKALEWVRQ
RAWKOLREAYSGDLOVPHRFQGVDSVYRRHAGNLETMRKGPYLVLLTPTAVKVE
GIPTWIIASHVVPKAPPDPSGWAEXTENPLRLLRHVVPYSVNNSS"
6185..8149
/gene="env"
6185..8149
/gene="env"
/codon_start=1
/product="envelope"
/protein_id="CAC39618.1"
/db_xref="GI:14275840"
/translation="MHPTLSRRHLPIRGKPRKIPLSFASIAWFLTLSTTPQVNGK
RLVDSPNHKLPLSLWLLTDSGTGININSTGEAPLGTPWPELYVCLRSVTPGLNDQA
TPDVLRYAGFYVPCPGPNNEEYCGNPQDFCKQWSCITSDNGWKNWPSQODRVSYS
FVNNPTSNOFNYGHRKQWQVORVKNQKQISCHSLDLDLYLKISFTEKQENIO
KWNQISGIVYGGSGRKKSVLIRLITOMEPVVAIGPNKGLAQGGPIQEQRP
SNPDSYNTSGSVTEPNITIKGAKLSLIQAFALNSTTEATSSCWLCNRTAE
PYTEGMAAGKFPNWKREHRDQCTWGSQNKLTLEVSGKGTGCMVPPSHOHLNRTAE
FNRTSELYLPGVDYRWACNTGLTPCVSTLVFNQTKDFCMVQIVPRVYVPEKAVL
DEQYLYRNPKRPEISLTSLSEVLQNRGLDLLFLKEGLGLVALKECCFCFVDSHGAI
DIOALEKSVNLNEESLTSLSEVLQNRGLDLLFLKEGLGLVALKECCFCFVDSHGAI
RDSMKLREKREKREKREKREKREKREKREKREKREKREKREKREKREKREKREK
LLNRVAVRQVSVNRVWLRLQVQGLPS"
8216..8918
/polya_signal
8810..8815
BASE COUNT 2435 a 2174 c 2253 g 2056 t
ORIGIN

```

```

Query Match 100.0%; Score 20; DB 14; Length 8918;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatggctctcctgaccttg 20
|||||
Db 5805 GATGGCTCTCTGCTCTTGT 5824

```

```

RESULT 8
BC005647 2435 bp mRNA ROD 12-JUL-2001
LOCUS Mus musculus, clone IMAGE:3709119, mRNA, partial cds.
DEFINITION BC005647
ACCESSION BC005647
VERSION BC005647.1 GI:13542908
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2687)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

## REMARK COMMENT

USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalob@bcm.tmc.edu](mailto:villalob@bcm.tmc.edu)  
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 17 Row: 1 Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

## FEATURES source

Location/Qualifiers  
1..2687  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3709119"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
ductal carcinoma, 5 month old virgin mouse."  
/clone\_lib="NCI\_CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
<1..2122  
/codon\_start=2  
/product="Unknown (protein for IMAGE:3709119)"  
/protein\_id="AAH05647.1"  
/db\_xref="GI:13542909"

## CDS

/translation="APVSVPCIPPHRLSKITNLTITQKSFQFMORLHSYVTLKROSR  
NGVPLRLTHLQSOBCEQVGRDSDKWKALKQKORLHDLERALLVELIR  
KREKLRETIKIQIQAEMQLTPELIRLRLTLEQLEKDTGNIFSEVPVLEVPDIL  
HTRKPDFFTKMKNLEAYRYLNFDDFEEDFNLVNCLKNYNAKDTIFYRAVRLRSG  
GAVLRQARQAKMGIDFETGHIHNLADGEVSHHTEDEVEERLLENKHLPVVE  
QLKLLLEDEVASQSVGRSRRAKMKIKETALRRKLAHRETGRDGPGRHPSGR  
GNLTPHACDKDQOTDSAAESSOETSKGLPNMSSTPAHEVGRRTSVLFSSKNPK  
TAGPKRPGRPKNRESQMPHSGSPVGPQPLIMGSLRQKRGSRPSSSSSDSDS  
KTETDPPMLPANGSSGNQPKKFLYVRCNLPSSDSSSSSSSSSSSSSSASDRT  
DTPSKQKGRKPSFSGTFPDSSEDTSTENAYSVGTGRGVGHSMVRKSLGRGAGW  
LSEDESDPLDLDLVAKRCGYSYPALIIDPKMPREGMHGVPVPPVPLEVILKE  
QMTQEAHEHLYLVLFNDKRTQWLPRTKLVPLGVNQDLDKEKMLEGRKSNTRKSVQI  
AYHRLQHRKSVQESSETSDSD"

BASE COUNT 717 a 733 c 740 g 497 t

Query Match 92.0%; Score 18.4; DB 10; Length 2687;  
Best Local Similarity 95.0%; Pred. No. 26;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctcctgaccttg 20

Db 2148 GATGGCTCTCTGCTCTTGT 2167

## RESULT 9

AX002804  
LOCUS AX002804 3482 bp DNA PAT  
DEFINITION Sequence 3 from Patent WO9853104.  
ACCESSION AX002804  
VERSION AX002804.1 GI:9885132  
KEYWORDS porcine endogenous retrovirus.  
SOURCE porcine endogenous retrovirus  
ORGANISM Viruses; Retroviridae; Mammalian type C

```

retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 3482)
Stoye,J.P. and Weiss,R.A.
Detection of retroviral subtypes based upon envelope specific
sequences
Patent: WO 9853104-A 3 26-NOV-1998;
MEDICAL RES COUNCIL (GB); STOYE JONATHAN PAUL (GB)
Location/Qualifiers
1. .3482
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 3482;
Best Local Similarity 95.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgcctttg 20
||| ||||||||||||||||
Db 528 GATAGCTCTCCTGCCCTTTG 547

RESULT 10
PERENV2 3482 bp RNA VRL 14-OCT-1997
DEFINITION Porcine endogenous retrovirus env gene, 3482 bp.
ACCESSION Y12239
VERSION Y12239.1 GI:2576324
KEYWORDS env gene.
SOURCE porcine endogenous retrovirus.
ORGANISM porcine endogenous retrovirus
Viruses: Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 3482)
Le Tissier,P., Stoye,J.P., Takeuchi,Y., Patience,C. and Weiss,R.A.
Two sets of human-tropic pig retrovirus
Nature 389 (6652), 681-682 (1997)
REFERENCE 97478526
AUTHORS Stoye,J.P.
TITLE Direct Submission
JOURNAL Submitted (02-APR-1997) J.P. Stoye, National Institute for Medical
Research, Virology, The Ridgeway Mill Hill, London, NW7 1AA, UK
REMARK revised by author 16-SEP-97
FEATURES
source
1. .3482
/organism="porcine endogenous retrovirus"
/specific_host="Sus scrofa"
/db_xref="taxon:61673"
/notes="PERV-B"
911..2884
/gene="env"
911..2884
/gene="env"
/codon_start=1
/protein_id="CAA72928.1"
/db_xref="GI:2576325"
/translation="MHPTLSWRHLPTRGEPKRLRPLSPASTAWFLTLTITPQASSK
RLIDSSNPHRPLSLTWLIIDPTGVTNTRGVAPRGTWMPHLFCLRLINPAVKSTP
PNLVRSYGFCCPTKEKVCYGGSGESFCRRSCVTSDGDKWPISLQDRVAFSPVN
SGPKYKVMKLYKDKSCPSDLDYKISTFGKQKQENLOKWINCMSWGIVFYKYGGA
GSTLILRLTETGTEPPVAVGPKVLAEGQPPALEPHNLPVQLTSLRPDITQPSN
GTTLITLPRNSPGVPVKTQGLFSLQGAFAQAINSTDPTDPTSSCLSLSSGPPY
EGMAKEGFWNTEKRNQCTGWSGRNKLTLEVSQKCTCKGAPPSHQLCYSTVVEQ
ASENQYLVPGYNRWACNTGLTFCVSTVFNQSKFCVMQIVPVRVYVHPVEVLDY
LYRNRKRPVSLTLAVMLGLTAVGVTGTAALITGPOOLEKGLGELHAAMTEDLR
ALFEVSNLSESLTSLSEVLQNRGLDLFLREGGLCAALKKECCFCYVDHSCAIRD
MSKRLRLRLRRERADQWFEFGFNRSFWMFTLLSALTGLPLVILLLLLTIVGFCLIN
RFVAFVRRVSAQIMVIRQOYQGLLSQGETDL"
BASE COUNT 927 a 854 c 867 g 834 t
ORIGIN

```

```

ORIGIN

Query Match 92.0%; Score 18.4; DB 14; Length 3482;
Best Local Similarity 95.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgcctttg 20
||| ||||||||||||||||
Db 528 GATAGCTCTCCTGCCCTTTG 547

RESULT 11
AX052635 4402 bp DNA PAT 12-JAN-2001
LOCUS Sequence 20 from Patent WO0071726.
DEFINITION AX052635
ACCESSION AX052635
VERSION AX052635.1 GI:12226825
KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
ORGANISM Viruses: Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4402)
Fiederspiel,M.J.
Methods to inhibit infectious agent transmission during
xenotransplantation
Patent: WO 0071726-A 20 30-NOV-2000;
Mayo Medical Ventures (US)
JOURNAL Location/Qualifiers
1. .4402
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1259 a 1085 c 1111 g 947 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 4402;
Best Local Similarity 95.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gatggctctcctgcctttg 20
||| ||||||||||||||||
Db 3301 GATAGCTCTCCTGCCCTTTG 3320

RESULT 12
AX052637 4918 bp DNA PAT 12-JAN-2001
LOCUS Sequence 22 from Patent WO0071726.
DEFINITION AX052637
ACCESSION AX052637
VERSION AX052637.1 GI:12226827
KEYWORDS porcine endogenous retrovirus.
SOURCE porcine endogenous retrovirus
ORGANISM Viruses: Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
1 (bases 1 to 4918)
Fiederspiel,M.J.
Methods to inhibit infectious agent transmission during
xenotransplantation
Patent: WO 0071726-A 22 30-NOV-2000;
Mayo Medical Ventures (US)
JOURNAL Location/Qualifiers
1. .4918
/organism="porcine endogenous retrovirus"
/db_xref="taxon:61673"
BASE COUNT 1382 a 1173 c 1183 g 1180 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 4918;
Best Local Similarity 95.0%; Pred. No. 27;

```

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20  
 Db 1804 GATAGCTCTCTGCCCTTG 1823

RESULT 13

AX052636 6076 bp DNA PAT 12-JAN-2001  
 LOCUS  
 DEFINITION Sequence 21 from Patent WO0071726.  
 AX052636  
 ACCESSION  
 VERSION AX052636.1 GI:12226826  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus

REFERENCE  
 AUTHORS Viruses; Retrovirdae; Retroviridae; Mammalian type C  
 TITLE retroviruses; 1-Mammalian type C virus group.  
 JOURNAL 1 (bases 1 to 6076)  
 FEATURES  
 source Methods to inhibit infectious agent transmission during  
 Location/Qualifiers  
 1. 6076

/organism="porcine endogenous retrovirus"

BASE COUNT 1613 a 1512 c 1609 g 1342 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 6076;  
 Best Local Similarity 95.0%; Pred. No. 28;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20  
 Db 5835 GATAGCTCTCTGCCCTTG 5854

RESULT 14  
 AX052647 6076 bp DNA PAT 12-JAN-2001  
 LOCUS  
 DEFINITION Sequence 32 from Patent WO0071726.  
 AX052647  
 ACCESSION  
 VERSION AX052647.1 GI:12226837  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus

REFERENCE  
 AUTHORS Viruses; Retrovirdae; Retroviridae; Mammalian type C  
 TITLE retroviruses; 1-Mammalian type C virus group.  
 JOURNAL 1 (bases 1 to 6076)  
 FEATURES  
 source Methods to inhibit infectious agent transmission during  
 Location/Qualifiers  
 1. 6076

/organism="porcine endogenous retrovirus"

BASE COUNT 1619 a 1505 c 1598 g 1354 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 6076;  
 Best Local Similarity 95.0%; Pred. No. 28;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20  
 Db 5835 GATAGCTCTCTGCCCTTG 5854

RESULT 15  
 AX052634 7362 bp DNA PAT 12-JAN-2001  
 LOCUS  
 DEFINITION Sequence 19 from Patent WO0071726.  
 AX052634  
 ACCESSION  
 VERSION AX052634.1 GI:12226824  
 KEYWORDS  
 SOURCE porcine endogenous retrovirus.  
 ORGANISM porcine endogenous retrovirus

REFERENCE  
 AUTHORS Viruses; Retrovirdae; Retroviridae; Mammalian type C  
 TITLE retroviruses; 1-Mammalian type C virus group.  
 JOURNAL 1 (bases 1 to 7362)  
 FEATURES  
 source Methods to inhibit infectious agent transmission during  
 Location/Qualifiers  
 1. 7362

/organism="porcine endogenous retrovirus"

BASE COUNT 1997 a 1821 c 1881 g 1663 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 6; Length 7362;  
 Best Local Similarity 95.0%; Pred. No. 28;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gatggctctctgcccttg 20  
 Db 5900 GATAGCTCTCTGCCCTTG 5919

Search completed: February 24, 2002, 02:56:21  
 Job time: 21292 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      February 24, 2002, 03:01:35 ; Search time 584.4 Seconds
              (without alignments)
              4870.499 Million cell updates/sec
```

Title:	US-09-171-553B-1
Perfect score:	3320

Sequence: 1 gaattcgcgccgctcgac.....aaaaaaaaaaaaaaaaaa 3320

Scoring table:

	OLIGO_NUC	Capent 60 0
Capent 60 0		

Searched: 930621 seqs, 428662619 residues

word size :

Total number of hits satisfying chosen parameters: 1861242

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

Database :

[illegible]

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3320	100.0	3320	18	AAV09698	Porcine retrovirus
2	2682	80.8	8309	18	AAV09700	Porcine retrovirus
3	2666	80.3	8196	18	AAV09699	Porcine retrovirus
4	1091	32.9	7333	22	AAE777823	Defective retrovirus
5	1091	32.9	7393	18	AA747883	Porcine retrovirus
6	848	25.5	4402	22	AAC67020	PERV env protein c
7	797	24.0	6076	22	AAC67021	PERV env protein c
8	797	24.0	7362	22	AAC67019	PERV env protein c
9	786	23.7	2462	20	AAV83748	Pig endogenous re
10	746	22.5	6076	22	AAC67032	PERV env protein c
11	704	21.2	7892	18	AA747884	Miniature swine r

12	704	11.2	8132	22	AAE77727	Nucleotide sequence
13	570	17.2	4918	22	AAE57022	PERV env protein c
14	498	15.0	600	22	AAH26252	pig endogenous ret
15	497	15.0	599	22	AAH26258	pig endogenous ret
16	497	15.0	599	22	AAH26260	pig endogenous ret
17	496	14.9	7873	22	AAE67023	PERV env protein c
18	474	14.3	1980	22	AAE67018	PERV-1-15 env prot
19	446	13.4	599	22	AAH26259	pig endogenous ret
20	438	13.2	540	22	AAH26253	pig endogenous ret
21	396	11.9	600	22	AAH26251	pig endogenous ret
22	396	11.9	600	22	AAH26254	pig endogenous ret
23	396	11.9	600	22	AAH26257	pig endogenous ret
24	395	11.9	599	22	AAH26255	pig endogenous ret
25	395	11.9	599	22	AAH26256	pig endogenous ret
26	376	11.3	478	22	AAH26283	pig endogenous ret
27	375	11.3	8060	18	AAAT74811	pig endogenous ret
28	375	11.3	8060	22	AAAT77125	Porcine retrovirus
29	356	10.7	458	22	AAH26264	Tsukuba-1 CDNA. p
30	325	9.8	478	22	AAH26265	pig endogenous ret
31	321	9.7	3482	20	AAVB2749	pig endogenous ret
32	301	9.1	478	22	AAH26266	pig endogenous ret
33	275	8.3	478	22	AAH26262	pig endogenous ret
34	229	6.9	458	22	AAH26261	pig endogenous ret
35	45	1.4	45	22	AAH26276	Human retrovirus D
36	41	1.2	45	22	AAH26275	Human retrovirus D
37	35	1.1	127	22	AAH26002	DoERV RT sequence,
38	35	1.1	8655	20	AAH269750	Mus dunali endogeno
39	32	1.0	463	22	AAH71398	Human cervical can
40	31	0.9	628	20	AAH207249	Mus telomerase R
41	31	0.9	955	22	AAE91898	Human secreted pro
42	31	0.9	955	22	AAE91925	Human secreted pro
43	31	0.9	1010	21	AAE18191	Lung cancer associ
44	31	0.9	1404	7	AAH260283	Sequence encoding
45	31	0.9	1404	8	AAH705005	Sequence encoding

	RESULT	1	
ID	AAV09698	standard; DNA; 3320 BP.	
XX			
AC	AAV09698;		
XX			
DT	19-MAY-1998	(first entry)	
XX			
DE	Porcine retrovirus pol and env DNA.		
XX			
KW	Porcine retrovirus; POBv; POL protein; ENV		
KW	diagnosis; xenotransplantation; prophylact		
XX			
OS	Porcine retrovirus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	23..2793	
FT		/tag= a	
FT		/product= POL protein	
FT		/note= "polymerase protein	
FT	CDS	2642..3297	
FT		/tag= b	
FT		/product= ENV protein	
FT		/note= "envelope protein"	
XX			
PN	MO9740167-A1.		
XX			
PD	30-OCT-1997.		
XX			
PF	18-APR-1997;	97MO-GBO1087.	
XX			
PR	10-FEB-1997;	97GB-0002668.	
PR	19-APR-1996;	96GB-0008164.	
XX			

PA (IMUT-) IMUTRAN LTD.  
PA (OONE-) Q-ONE BIOTECH LTD.

PI Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX Polynucleotide encoding porcine retrovirus expression product -  
PT useful to develop products for use in vaccines, diagnosis and  
PT xeno-transplantation

XX Claim 4: Fig 1: 69pp; English.

XX This DNA sequence encodes the porcine retrovirus (PoEV) polymerase (POL)  
CC and envelope (ENV) proteins. These proteins can be used to develop viral  
CC vaccines, antisense nucleic acids, ribozymes and other antiviral agents.  
CC They can also be used in xeno-transplantation technology and as  
CC diagnostic tools.

XX Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;

Query Match 100.0%; Score 3320; DB 18; Length 3320;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcgccgagcgttcgaacagatgctcttctcgtctgagatatacccccactagcaaa 60  
DB 1 gaattcgccgagcgttcgaacagatgctcttctcgtctgagatatacccccactagcaaa 60  
QY 61 ccaatttttgccttcgaatggaagatccagatcaggaagaacccgagcagctcaactg 120  
DB 61 ccaatttttgccttcgaatggaagatccagatcaggaagaacccgagcagctcaactg 120  
QY 121 acccgactgccccaaaggttcaaaaactcccccacattttgaacgaagcccttaacagg 180  
DB 121 acccgactgccccaaaggttcaaaaactcccccacattttgaacgaagcccttaacagg 180  
QY 121 acccgactgccccaaaggttcaaaaactcccccacattttgaacgaagcccttaacagg 180  
DB 121 acccgactgccccaaaggttcaaaaactcccccacattttgaacgaagcccttaacagg 180  
QY 181 gacctggccaactcagatccaaacacccctcagtgacctccctcagtaagtgatgac 240  
DB 181 gacctggccaactcagatccaaacacccctcagtgacctccctcagtaagtgatgac 240  
QY 241 ctgctcttcgagcagcccaacacagactctgttaagaagtcagaaagcactactctg 300  
DB 241 ctgctcttcgagcagcccaacacagactctgttaagaagtcagaaagcactactctg 300  
QY 301 gaattgtctgacctaggttaacagcctctgttaagaagtcagaaagcactactctg 360  
DB 301 gaattgtctgacctaggttaacagcctctgttaagaagtcagaaagcactactctg 360  
QY 361 gtaacatacttcgggttaacagcttccggggcgagcgatggtcgagagcgagcgagaa 420  
DB 361 gtaacatacttcgggttaacagcttccggggcgagcgatggtcgagagcgagcgagaa 420  
QY 421 aaaactgtatccagatccgcccacacacacagcacaagaatgagagattttgggg 480  
DB 421 aaaactgtatccagatccgcccacacacacagcacaagaatgagagattttttgggg 480  
QY 481 acaagctgatttcagagactgtagatcccgaggttttcgaccttaagagcccccctac 540  
DB 481 acaagctgatttcagagactgtagatcccgaggttttcgaccttaagagcccccctac 540  
QY 541 acagctgatttcagagactgtagatcccgaggttttcgaccttaagagcccccctac 540  
DB 541 acagctgatttcagagactgtagatcccgaggttttcgaccttaagagcccccctac 540  
QY 541 ccgcttaacaaagaaagggggaattccctgggctccttcgagcaccgaaggaatttgat 600  
DB 541 ccgcttaacaaagaaagggggaattccctgggctccttcgagcaccgaaggaatttgat 600  
QY 601 gctacaaagagccctctgagcgacactgctctggccctcctgacgtactaaacc 660  
DB 601 gctacaaagagccctctgagcgacactgctctggccctcctgacgtactaaacc 660  
QY 661 ttacccttattgtgagagcgtaaggaagtaagccgagagagtttaacccaacta 720  
DB 661 ttacccttattgtgagagcgtaaggaagtaagccgagagagtttaacccaacta 720

QY 721 ggacattgagagagactctgttcctactgtcaaaagacttgatccctgtagcagtggt 780  
DB 721 ggacattgagagagactctgttcctactgtcaaaagacttgatccctgtagcagtggt 780  
QY 781 tggcccgatgtctgaagactatcgacgtctgtgacactggttcaagaagcgtgacaaa 840  
DB 781 tggcccgatgtctgaagactatcgacgtctgtgacactggttcaagaagcgtgacaaa 840  
QY 841 ttgacttggagcaataataactgttaataagcccccatgcatctggagaacatctgcg 900  
DB 841 ttgacttggagcaataataactgttaataagcccccatgcatctggagaacatctgcg 900  
QY 901 cagccccagacagatgtagtgaacacagcccgatagacccaatcaaaagcctgtctc 960  
DB 901 cagccccagacagatgtagtgaacacagcccgatagacccaatcaaaagcctgtctc 960  
QY 961 acagagagaggtactctgcctccacagcgtctccaaactgacactctctgctgaa 1020  
DB 961 acagagagaggtactctgcctccacagcgtctccaaactgacactctctgctgaa 1020  
QY 1021 ggaactgtgaacacagtgatgatgtagcctcaactatgattgagagagctgggtc 1080  
DB 1021 ggaactgtgaacacagtgatgatgtagcctcaactatgattgagagagctgggtc 1080  
QY 1081 cgaagagacacttaacagacataccgctgtagagagtgtaacctgttcaactgacgga 1140  
DB 1081 cgaagagacacttaacagacataccgctgtagagagtgtaacctgttcaactgacgga 1140  
QY 1141 agcagctatgtgtgtaagagtaagagatggttggtggcgaggtgtggaacggagccgc 1200  
DB 1141 agcagctatgtgtgtaagagtaagagatggttggtggcgaggtgtggaacggagccgc 1200  
QY 1201 acgatctggcagcagcagccttcgcggaaggaactcagcgcgaagagctgagctcagc 1260  
DB 1201 acgatctggcagcagcagccttcgcggaaggaactcagcgcgaagagctgagctcagc 1260  
QY 1261 ctcaagcaagcttgcgctgcgaggaaggaatccataaatcttaacgagcagcg 1320  
DB 1261 ctcaagcaagcttgcgctgcgaggaaggaatccataaatcttaacgagcagcg 1320  
QY 1321 tatgcttctgagctgacagctgacagctgacagctgacagctgacagctgacagctg 1380  
DB 1321 tatgcttctgagctgacagctgacagctgacagctgacagctgacagctgacagctg 1380  
QY 1381 tcagcagagaggaataaagaacaaagagaaatcttaagctatlaagaagccttaacat 1440  
DB 1381 tcagcagagaggaataaagaacaaagagaaatcttaagctatlaagaagccttaacat 1440  
QY 1441 ttgcaaaaagagctagctatlaacactgtccttgagacatcagaagaagccaagatctcata 1500  
DB 1441 ttgcaaaaagagctagctatlaacactgtccttgagacatcagaagaagccaagatctcata 1500  
QY 1501 tctagaaggaacacagatggtctgacgggtttgccaagcagcgccagagctgttaacct 1560  
DB 1501 tctagaaggaacacagatggtctgacgggtttgccaagcagcgccagagctgttaacct 1560  
QY 1561 ctgctcataatagaacacgcccagaagcccaagcccaagcccaagcagagcagcctagaagac 1620  
DB 1561 ctgctcataatagaacacgcccagaagcccaagcccaagcccaagcagagcagcctagaagac 1620  
QY 1621 ttgcaagagataaaagaatagacagctctctgagactccggaagggcgctgtataac 1680  
DB 1621 ttgcaagagataaaagaatagacagctctctgagactccggaagggcgctgtataac 1680  
QY 1681 tcatatggagagaaatccctgcccagaagaggtttgatatgttccaacagatatac 1740  
DB 1681 tcatatggagagaaatccctgcccagaagaggtttgatatgttccaacagatatac 1740  
QY 1741 cgtctaacaccactagagaaataacacctgacagcaatttggtcagaacatcccttatac 1800  
DB 1741 cgtctaacaccactagagaaataacacctgacagcaatttggtcagaacatcccttatac 1800



QY 1801 gtctgaggtacacagagtggtgactcgggtgltcaaacattgtgtccctgccaagt 1860  
 Db 1801 gtctgaggtacacagagtggtgactcgggtgltcaaacattgtgtccctgccaagt 1860  
 QY 1861 gttatgttaattcttccagataatcctccaggaagaagactaaggggaagccaccagc 1920  
 Db 1861 gttatgttaattcttccagataatcctccaggaagaagactaaggggaagccaccagc 1920  
 QY 1921 gctactggaagtggaactcactgaagtgaaagcgggtactaatccggaacaaatatacta 1980  
 Db 1921 gctactggaagtggaactcactgaagtgaaagcgggtactaatccggaacaaatatacta 1980  
 QY 1981 ttggtttttgtagacacctttccagagtgagggtctatccctactaagaagaagact 2040  
 Db 1981 ttggtttttgtagacacctttccagagtgagggtctatccctactaagaagaagact 2040  
 QY 2041 tcaacctgtgtggttaagaaatactctggaagaaattttccaaagatttggaaatacctaag 2100  
 Db 2041 tcaacctgtgtggttaagaaatactctggaagaaattttccaaagatttggaaatacctaag 2100  
 QY 2101 gtaataagggtcaagcaaatgttccagcttctgtgcccagtgaaagtcaggagtcgccaag 2160  
 Db 2101 gtaataagggtcaagcaaatgttccagcttctgtgcccagtgaaagtcaggagtcgccaag 2160  
 QY 2161 atattggggatltgacttggaactgtgcatltgcatatacagcccaagctcaggacagta 2220  
 Db 2161 atattggggatltgacttggaactgtgcatltgcatatacagcccaagctcaggacagta 2220  
 QY 2221 gaggaggtgaataagacatlaagagaccccttaacaaattgaccacaaagactgtgcat 2280  
 Db 2221 gaggaggtgaataagacatlaagagaccccttaacaaattgaccacaaagactgtgcat 2280  
 QY 2281 aatgattggtatgtgtccctgtccctgtgtctttagagggtgaggaacaccccttggaag 2340  
 Db 2281 aatgattggtatgtgtccctgtccctgtgtctttagagggtgaggaacaccccttggaag 2340  
 QY 2341 ttgggtgtgacccctctatgtaattgtctacggggagaccccccgttggcagaatgtgc 2400  
 Db 2341 ttgggtgtgacccctctatgtaattgtctacggggagaccccccgttggcagaatgtgc 2400  
 QY 2401 ttgacatatgtgtagt 2460  
 Db 2401 ttgacatatgtgtagt 2460  
 QY 2461 gagggtgtagcagcagcagcgttggaagcagctccggagagcctactcagaggaagactgt 2520  
 Db 2461 gagggtgtagcagcagcagcgttggaagcagctccggagagcctactcagaggaagactgt 2520  
 QY 2521 caagttccacatcgtcttccaaagtgtgagatcagctatgtagttagacgcgcagtcagaa 2580  
 Db 2521 caagttccacatcgtcttccaaagtgtgagatcagctatgtagttagacgcgcagtcagaa 2580  
 QY 2581 aacctcgaagctcgttggaagggacactatctgtactttttagacacaaagcgtgtgt 2640  
 Db 2581 aacctcgaagctcgttggaagggacactatctgtactttttagacacaaagcgtgtgt 2640  
 QY 2641 aaagtgcgaaggaatcccccacttgatccatgcaatcccaagtgaaagcgcgcagtcgt 2700  
 Db 2641 aaagtgcgaaggaatcccccacttgatccatgcaatcccaagtgaaagcgcgcagtcgt 2700  
 QY 2701 gattcgggtgtggaagc 2760  
 Db 2701 gattcgggtgtggaagc 2760  
 QY 2761 gtctctactctgtcaataactcctcaagttaattgtgtaagcgcgtgtggaagcccgaa 2820  
 Db 2761 gtctctactctgtcaataactcctcaagttaattgtgtaagcgcgtgtggaagcccgaa 2820  
 QY 2821 ctccataaacccttatctcactcgtgtactactactactactactactactactactactat 2880  
 Db 2821 ctccataaacccttatctcactcgtgtactactactactactactactactactactactat 2880  
 QY 2881 taacagcactcaaggggaggtccctgtgggagctgtgtgcccgaattatagtcgtcgt 2940

Db 2881 taacagcactcaaggggaggtccctgtgggagctgtgtgcccgaattatagtcgtcgt 2940  
 QY 2941 tggataagtaattcctcgtgtctcaatgacacagccacaccccgatgtactcgtgtctta 3000  
 Db 2941 tggataagtaattcctcgtgtctcaatgacacagccacaccccgatgtactcgtgtctta 3000  
 QY 3001 cgggttttcgtgttgc 3060  
 Db 3001 cgggttttcgtgttgc 3060  
 QY 3061 ttctccttgcaagcaatgagctgcaataacttcaatgtagtgggaattggaatgccaagt 3120  
 Db 3061 ttctccttgcaagcaatgagctgcaataacttcaatgtagtgggaattggaatgccaagt 3120  
 QY 3121 ctctcagcaagaagagtagtactcctttgttaacactcctaccagttataatcaat 3180  
 Db 3121 ctctcagcaagaagagtagtactcctttgttaacactcctaccagttataatcaat 3180  
 QY 3181 taattatggccatggaagatggaagattggaagacagcggtacaaagaatgtacgaa 3240  
 Db 3181 taattatggccatggaagatggaagattggaagacagcggtacaaagaatgtacgaa 3240  
 QY 3241 taagcaataaagctgtcattcgtttagacactagattacttaaaataaagttactactaaa 3300  
 Db 3241 taagcaataaagctgtcattcgtttagacactagattacttaaaataaagttactactaaa 3300  
 QY 3301 aaaaaaaaaaaaaaaaaa 3320  
 Db 3301 aaaaaaaaaaaaaaaaaa 3320

RESULT 2  
 AAV09700  
 ID AAV09700 standard; DNA; 8209 BP.  
 XX  
 XX AAV09700;  
 DT 20-MAY-1998 (first entry)  
 DE Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.  
 XX  
 XX Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;  
 KW virion core polypeptide; polymerase protein; envelope protein;  
 KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
 XX  
 OS Porcine retrovirus.  
 FH  
 FH Key Location/Qualifiers  
 FT 1..61  
 FT misc\_feature  
 FT /tag= a  
 FT /note= "R LTR domain"  
 FT 1..588  
 FT LTR  
 FT /tag= b  
 FT 62..143  
 FT misc\_feature  
 FT /tag= c  
 FT /note= "u5 LTR domain"  
 FT 588..2162  
 FT CDS  
 FT /tag= d  
 FT /product= GAG protein  
 FT /note= "viral core polypeptide"  
 FT 2163..5747  
 FT CDS  
 FT /tag= e  
 FT /product= POL protein  
 FT /note= "polymerase peptide"  
 FT 5748..7590  
 FT CDS  
 FT /tag= f  
 FT /product= ENV protein  
 FT /note= "envelope protein"  
 FT 7591..8209  
 FT LTR  
 FT /tag= g  
 FT 7638..8106  
 FT misc\_feature  
 FT /tag= h

FT	protein_bind	/note="U3 LTR domain"
FT		7643..7686
FT		/*tag= i
FT		/note= "Binding site for transcription factors
FT		OCT-1, c-Myb, Lf-1, E47 and ETS-1"
FT	protein_bind	7719..7739
FT		/*tag= j
FT		/note= "Binding site for transcription factors
FT		AP-4 and NF1"
FT	protein_bind	7744..7754
FT		/*tag= k
FT		/note= "Binding site for transcription factor AP-1/TR"
FT		7773..7785
FT		/*tag= l
FT		/note= "Binding site for transcription factor ETS-1/GATA"
FT		7793..7884
FT		/*tag= m
FT		/note= "Binding site for transcription factors
FT		GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT	protein_bind	7841..7851
FT		/*tag= n
FT		/note= "Binding site for transcription factor AP-1"
FT		7873..7883
FT	protein_bind	7885..7914
FT		/*tag= o
FT		/note= "Binding site for transcription factor AP-1"
FT		7885..7914
FT		/*tag= p
FT		/note= "Binding site for transcription factors
FT		AP-1/GATA, c-Myb and AP-1"
FT	protein_bind	7928..7953
FT		/*tag= q
FT		/note= "Binding site for transcription factors E47
FT		and ETS-1"
FT	CANT_signal	7964..7968
FT		/*tag= r
FT		7985..7997
FT	protein_bind	/*tag= s
FT		/note= "Binding site for transcription factor GATA"
FT		8003..8007
FT	CANT_signal	8024..8036
FT		/*tag= t
FT		8042..8046
FT	protein_bind	/*tag= u
FT		/note= "Binding site for transcription factor GATA"
FT		8042..8046
FT		/*tag= v
FT		8049..8060
FT		/*tag= w
FT		/note= "Binding site for transcription factor
FT		AP-1/CREB"
FT	TATA_signal	8075..8090
FT		/*tag= x
FT		1..61
FT	misc_feature	/*tag= y
FT		/note= "R LTR domain"
FT		8127..8135
FT	misc_feature	/*tag= z
FT		/note= "PolyA downstream element"
FT		8161..8166
FT	polyA_signal	/*tag= aa
FT		
XX	W09740167-A1.	
XX		
XX	30-OCT-1997.	
XX		
XX	18-APR-1997;	97MO-GB01087.
XX		
XX	10-FEB-1997;	97GB-0002668.
XX	19-APR-1996;	96GB-0008164.
XX		
XX	(IMUT-) IMUTRAN LTD.	
XX	(OONE-) O-ONE BIOTECH LTD.	
XX		
XX	Galbraith DN, Haworth C, Lees GM, Smith KT;	

[illegible]

Db	3692	tgccataccgtgtaagaagcttgatctccttgagccaggtgtggccgtatgctcgaagc	3751
QY	801	tatcgacgtgtggtgacatctggtcgaagagcgttgacaatttgactttggagcagaatat	860
Db	3752	tatcgacgtgtggtgacatctggtcgaagagcgttgacaatttgactttggagcagaatat	3811
QY	861	aactgttaagcccccatgcatctggagaacatcgttcgcgaagcccccgagccgattgat	920
Db	3812	aactgttaagcccccatgcatctggagaacatcgttcgcgaagcccccgagccgattgat	3871
QY	921	gaaccaagcccgatgacccaatacaagaagctgtcttcaagaagaggttcactctgcg	980
Db	3872	gaaccaagcccgatgacccaatacaagaagctgtcttcaagaagaggttcactctgcg	3931
QY	981	tccacaagccgcttcacaaccttcgactcttcgccttggaagagcatgatgaacccgtac	1040
Db	3932	tccacaagccgcttcacaaccttcgactcttcgccttggaagagcatgatgaacccgtac	3991
QY	1041	tcatgattgcatacaactatgattgtaggaagacttggtgtcgcgaagagacttacagacat	1100
Db	3992	tcatgattgcatacaactatgattgtaggaagacttggtgtcgcgaagagacttacagacat	4051
QY	1101	accgcttgacttggaagaatgtgtaaaccttggttcacttgacggaagcagctatgtgtggaag	1160
Db	4052	accgcttgacttggaagaatgtgtaaaccttggttcacttgacggaagcagctatgtgtggaag	4111
QY	1161	taagagagctggtcggtgcgcagtggtgtgtagcagagcccaagatctgtggccaagacct	1220
Db	4112	taagagagctggtcggtgcgcagtggtgtgtagcagagcccaagatctgtggccaagacct	4171
QY	1221	gccggaaggaacttcagcgcgaagaagcttgagctcatggtccctcaccgcaagctttgcgct	1280
Db	4172	gccggaaggaacttcagcgcgaagaagcttgagctcatggtccctcaccgcaagctttgcgct	4231
QY	1281	ggccgaaggggaatccataaacattatacgaacagcagtagtgctttggcagctgcaca	1340
Db	4232	ggccgaaggggaatccataaacattatacgaacagcagtagtgctttggcagctgcaca	4291
QY	1341	cgtaacaggggcatctataacaagaagggtgtgtcttaccctcagcagggaggaaataa	1400
Db	4292	cgtaacaggggcatctataacaagaagggtgtgtcttaccctcagcagggaggaaataa	4351
QY	1401	gaacaaagaggaataatttaagcctattagaagccttacattggccaanaagctagcat	1460
Db	4352	gaacaaagaggaataatttaagcctattagaagccttacattggccaanaagctagcat	4411
QY	1461	tatacactgtcttcgacatcagaagaagccaaagatcctcatcttlaagaaggaccagatgac	1520
Db	4412	tatacactgtcttcgacatcagaagaagccaaagatcctcatcttlaagaaggaccagatgac	4471
QY	1521	tgaccgggttgccaagaagcagccccaagcgtgttaaccttctgcctataatagaacagcc	1580
Db	4472	tgaccgggttgccaagaagcagccccaagcgtgttaaccttctgcctataatagaacagcc	4531
QY	1581	caaaagcccaagaccocccagacagcagtagaccccttagaagatgycgaagagataaaaaagat	1640
Db	4532	caaaagcccaagaccocccagacagcagtagaccccttagaagatgycgaagagataaaaaagat	4591
QY	1641	agaccagttctcttgagactccggaggggacctgtcataccttaatgtyggaaggaaacct	1700
Db	4592	agaccagttctcttgagactccggaggggacctgtcataccttaatgtyggaaggaaacct	4651
QY	1701	ggcccaacaagaagggttgagaatgttccaagaagataatgtctcaacccacctaagaac	1760
Db	4652	ggcccaacaagaagggttgagaatgttccaagaagataatgtctcaacccacctaagaac	4711
QY	1761	taaacaccttgacagcttggttcagaacaatccacctatacatgttctggaagctacaagagat	1820
Db	4712	taaacaccttgacagcttggttcagaacaatccacctatacatgttctggaagctacaagagat	4771
QY	1821	ggcgtactggtgtgtaaacattgtgtgcccgcgaagctgtgtatgctaactctccag	1880

Dh	4772	ggtcgactccggttggtaacaacattgtgtgctccctccgactcgtgttaatgcttaactcctccaag	4831
QY	1861	aatactctccaggaagaagactaaaggggaagccaccccaagcgctcactctggaagctgagactt	1940
Dh	4832	aatactctccaggaagaagactaaaggggaagccaccccaagcgctcactctggaagctgagactt	4831
QY	1941	cactgaagtaagaagccgacttaataacggaacaaatactatgtgtgtttgttagaacactt	2000
Dh	4892	cactgaagtaagaagccgacttaataacggaacaaatactatgtgtgtttgttagaacactt	4951
QY	2001	ttcaggaatggtaagagacttatacttaagaagaagagacttcaacgctgtgtgcttaagaa	2060
Dh	4952	ttcaggaatggtaagagacttatacttaagaagaagagacttcaacgctgtgtgcttaagaa	5011
QY	2061	aatactgaaggaataatttttccaagaattttggaataacttaagatgaataggttcaagaatg	2120
Dh	5012	aatactgaaggaataatttttccaagaattttggaataacttaagatgaataggttcaagaatg	5071
QY	2121	ttcagactctgtcttccccagtaagtcagagagactgccccaaagatatgtggagattgattgaa	2180
Dh	5072	ttcagactctgtcttccccagtaagtcagagagactgccccaaagatatgtggagattgattgaa	5131
QY	2181	actgcattgtgcataaagaccccccaagactcaaggaacaggttagagagatgatataagaacat	2240
Dh	5132	actgcattgtgcataaagaccccccaagactcaaggaacaggttagagagatgatataagaacat	5191
QY	2241	taagaagagaccccttaacaaatttgacacaaagagactgtgcatatagtttgattgagctctct	2300
Dh	5192	taagaagagaccccttaacaaatttgacacaaagagactgtgcatatagtttgattgagctctct	5251
QY	2301	gaccttgtgcttttttagaggtagaagacaccccttgagacagtttggctgcagccccctatga	2360
Dh	5252	gaccttgtgcttttttagaggtagaagacaccccttgagacagtttggctgcagccccctatga	5311
QY	2361	attgctctaaggaggagacccccccgctgtggcagaataattgctcttgcaactagtctgatatgt	2420
Dh	5312	attgctctaaggaggagacccccccgctgtggcagaataattgctcttgcaactagtctgatatgt	5371
QY	2421	gctgcttctccaaagccttgtctctcttagagtcctaaagcgctgcagatgtgtgtgtagagcagagc	2480
Dh	5372	gctgcttctccaaagccttgtctctcttagagtcctaaagcgctgcagatgtgtgtgtagagcagagc	5431
QY	2481	gttgaaagcaactcccggaagagcctcactcccaaggaagagacttgcaagttccacactgcttcca	2540
Dh	5432	gttgaaagcaactcccggaagagcctcactcccaaggaagagacttgcaagttccacactgcttcca	5491
QY	2541	agttggagatctcaagtcatagtttgaacgacacacgttcagaagaacctcagagactcgtgtgaa	2600
Dh	5492	agttggagatctcaagtcatagtttgaacgacacacgttcagaagaacctcagagactcgtgtgaa	5551
QY	2601	ggagacctaactcgtacttlttgacacacacacacgagctgttgaaagtgcgaaggaatccccac	2660
Dh	5552	ggagacctaactcgtacttlttgacacacacacacgagctgttgaaagtgcgaaggaatccccac	5611
QY	2661	ctggatccatgatccacacgattaaagccggcgccactccga	2702
Dh	5612	ctggatccatgatccacacgattaaagccggcgccactccga	5653
RESULT 3			
AAV09699			
ID	AAV09699 standard; DNA; 8136 BP.		
AC	AAV09699;		
XX	19-MAY-1998 (first entry)		
XX	Porcine retrovirus DNA encoding, GAG, POL and ENV.		
DE	Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;		
KM	vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds		
XX	Porcine retrovirus.		
XX			

XX	Key	Location/Qualifiers
FH	CDS	576..2126
FT		/*tag= a
FT		/product= GAG protein
FT		/note= "villon core polypeptide"
FT		2143..5733
FT	CDS	/*tag= b
FT		/product= POL protein
FT		/note= "polymerase peptide sequence as given in specification"
FT		5606..7576
FT	CDS	/*tag= c
FT		/product= ENV protein
FT		/note= "envelope protein"
XX		
PX		W09740167-A1.
PD		30-OCT-1997.
XX		
PF		18-APR-1997; 97WO-GB01087.
XX		
PR		10-FEB-1997; 97GB-0002668.
PR		19-APR-1996; 96GB-0008164.
XX		(IMUT-) IMUTRAN LTD.
PA		(OONE-) O-ONE BIOTECH LTD.
PI		Gabraith DN, Haworth C, Lees GM, Smith KT;
XX		WPI: 1997-535851/49.
XX		
PT		Polynucleotide encoding porcine retrovirus expression product -
PT		useful to develop products for use in vaccines, diagnosis and
PT		xeno-transplantation
XX		
PS		Claim 4; Fig 2; 6pp: English.
XX		
CC		This DNA sequence encodes the porcine retrovirus (POEV) villon core
CC		polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC		proteins can be used to develop viral vaccines, antisense nucleic acids,
CC		ribozymes and other antiviral agents. They can also be used in
CC		xeno-transplantation technology and as diagnostic tools.
XX		
SQ		Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

Query Match            80.3%; Score 2666; DB 18; Length 8196;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      21  agatgccttcttcgtcgtgagattacaccccccactggccaaccttttgcttgaatg  80
DB      2958  aaatgccttcttcgtcgtgagattacacccccactggccaaccttttgcttgaatg  3017
QY      81  gggagatccaggtaacggagaaccggcgagctcaccttgagaccgactgcccaagggt  140
DB      3018  gggagatccaggtaacggagaaccggcgagctcaccttgagaccgactgcccaagggt  3077
QY      141  caagaactcccgacaattttgaagaaagccctaacacagagactggccaattcaagt  200
DB      3078  caagaactcccgacaattttgaagaaagccctaacacagagactggccaattcaagt  3137
QY      201  caaacacccctcagggtaacctctcctccagtaagtgtaaccttgcttgcggaagccac  260
DB      3138  ccaaacacccctcagggtaacctctcctccagtaagtgtaaccttgcttgcggaagccac  3197
QY      261  caaacagagactcttagaaggtacgaaggcactactgctggsaatgtctgcactaggcta  320
DB      3198  caaacagagactcttagaaggtacgaaggcactactgctggsaatgtctgcactaggcta  3257
QY      321  cagagccttgtcagaagaagccagaattggcgggagaaggyfnaactattlgggtacag  380
          |||||||
    
```

[illegible]

```

QY 1461 tatacactgtccttgagatcagaaagccaagatcatalctagagggaaccagatgac 1520
    |||||
Db 4398 tatacactgtccttgagatcagaaagccaagatcatalctagagggaaccagatgac 4457
QY 1521 tgaccgggttcaggaagcagccagagctgttaaccttcgtcctaataagaaagcc 1560
    |||||
Db 4458 tgaccgggttcaggaagcagccagagctgttaaccttcgtcctaataagaaagcc 4517
QY 1581 caaagcccaagaaagcagcagctacaccttagaagactggaagaagataaagaat 1640
    |||||
Db 4518 caaagcccaagaaagcagcagctacaccttagaagactggaagaagataaagaat 4577
QY 1641 agaccagttctctgagactccggaggagctgatacctcatalctagagggaagaaacct 1700
    |||||
Db 4578 agaccagttctctgagactccggaggagctgatacctcatalctagagggaagaaacct 4637
QY 1701 gccccacaagaagaaggttagaataatgtccacaagatacatcgtctaaaccactagaac 1760
    |||||
Db 4638 gccccacaagaagaaggttagaataatgtccacaagatacatcgtctaaaccactagaac 4697
QY 1761 taacacactgagcaggttggtcacaacatccctcatalctagagggttaccagagc 1820
    |||||
Db 4698 taacacactgagcaggttggtcacaacatccctcatalctagagggttaccagagc 4757
QY 1821 ggcctgactcgggtggtcacaacatgtgtccctcgagctggttaatactcctccag 1880
    |||||
Db 4758 ggcctgactcgggtggtcacaacatgtgtccctcgagctggttaatactcctccag 4817
QY 1881 aatacctccagggaagagactaaggggaagccacccagcgctcactggaagtgacct 1940
    |||||
Db 4818 aatacctccagggaagagactaaggggaagccacccagcgctcactggaagtgacct 4877
QY 1941 cactgaggtaaagccggctaaatacaggaacaatatctgtgtttttagaacctt 2000
    |||||
Db 4878 cactgaggtaaagccggctaaatacaggaacaatatctgtgtttttagaacctt 4937
QY 2001 ttcaagatggtagaggtctatcctactaagaaagagactccaacgtgtgtcagaag 2060
    |||||
Db 4938 ttcaagatggtagaggtctatcctactaagaaagagactccaacgtgtgtcagaag 4997
QY 2061 aatactggaagaaattttccaagatttgaataactaagatagtggtcagaagatg 2120
    |||||
Db 4998 aatactggaagaaattttccaagatttgaataactaagatagtggtcagaagatg 5057
QY 2121 tccagcttcgttgcacaggttaagtcagggactggccaagatatgtgggatgtatgga 2180
    |||||
Db 5058 tccagcttcgttgcacaggttaagtcagggactggccaagatatgtgggatgtatgga 5117
QY 2181 actgcatgtgcatlacagaccacaagctcagagcagtagagagatgaataacat 2240
    |||||
Db 5118 actgcatgtgcatlacagaccacaagctcagagcagtagagagatgaataacat 5177
QY 2241 taaagagacccttaccacaatattgaccacagagactggtcatalatgtatggtctcct 2300
    |||||
Db 5178 taaagagacccttaccacaatattgaccacagagactggtcatalatgtatggtctcct 5237
QY 2301 gcccttgtgcttttagaggtgaggaacaccccttgacaagtttgggtcgaaccccatba 2360
    |||||
Db 5238 gcccttgtgcttttagaggtgaggaacaccccttgacaagtttgggtcgaaccccatba 5297
QY 2361 attgctctacaggggaaccccccgttgcagaaattgtccttgcacatagtgatgtc 2420
    |||||
Db 5298 attgctctacaggggaaccccccgttgcagaaattgtccttgcacatagtgatgtc 5357
QY 2421 gctgcttcccaagccttttctctaggtcgaagcgctcaggtggtgagcagcgagc 2480
    |||||
Db 5358 gctgcttcccaagccttttctctaggtcgaagcgctcaggtggtgagcagcgagc 5417
QY 2481 gtaggaagcagctcccggaagcctactcagagagagagacttgcagattccacatcgcttca 2540
    |||||
Db 5418 gtaggaagcagctcccggaagcctactcagagagagagacttgcagattccacatcgcttca 5477

```

```

QY 2541 agltgagatcagctatgtatgtagaccacacgctgaggaagaaactgagactcgttggaa 2600
    |||||
Db 5478 agltgagatcagctatgtatgtagaccacacgctgaggaagaaactgagactcgttggaa 5537
QY 2601 gggacctatctcgtacttttggacacacaaagcgtgtgaaagtgcagaagatcccccac 2660
    |||||
Db 5538 gggacctatctcgtacttttggacacacaaagcgtgtgaaagtgcagaagatcccccac 5597
QY 2661 ctgtatccatgcatcccaaggttaacg 2686
    |||||
Db 5598 ctgtatccatgcatcccaaggttaacg 5623

RESULT 4
AAF77726
ID AAF77726 standard, DNA; 7333 BP.
XX
AC AAF77726;
XX
DT 23-MAY-2001 (first entry)
XX
DE Defective retroviral genome isolated from PK-15 cell line.
XX
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
ss.
XX
OS Unidentified.
XX
PN US6190861-B1.
XX
PD 20-FEB-2001.
XX
PF 13-DEC-1996; 96US-0766528.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA.
XX
DR WPI: 2001-256211/26.
XX
DR P-PSDB; AAB73282, AAB73283, AAB73284.
XX
PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome.
XX
PS Claim 1; Fig 2; 127pp; English.
XX
CC The present invention relates to a method for screening a cell or tissue
CC for the presence or expression of a retrovirus (RV), comprising
CC contacting a target nucleic acid from the cell or tissue with a second
CC nucleic acid from the present invention (e.g. the present sequence or a
CC fragment thereof). The method is useful for RV detection and to assess
CC graft transplantation risk. Screening of animals allows the elimination
CC of donors with active replication of known viruses. Inactive viruses
CC can be detected and inactivated, allowing identification and elimination
CC of potential human pathogens derived from swine in a manner not possible
CC in the outbred human organ donor population and is important to the
CC development of human xenotransplantation.
XX
SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 32.9%; Score 1091; DB 22; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1364 aaagggggttgccttaacctcagcagggagggaataaagaagaagaattcttaagcc 1423
    |||||
Db 3517 aaagggggttgccttaacctcagcagggagggaataaagaagaagaattcttaagcc 3576
QY 1424 tatlagagccttacatttgcacaaagcagctagatatatacactgtccctgagatcaga 1483

```

```

|||||
Db 3577 tattagaagccttaccatttggccaaaaggctagctatataacatctctctgacatcaga 3636
QY 1484 aagccaaagatctcatatcttagagggaaccagatgctacccgggtgtccaaagcagcg 1543
Db 3637 aagccaaagatctcatatcttagagggaaccagatgctacccgggtgtccaaagcagcg 3696
QY 1544 cccagctgttaacctctctgctatataatagaacgcccaagccccaagaccagaagcagc 1603
Db 3697 cccagctgttaacctctctgctatataatagaacgcccaagccccaagaccagaagcagc 3756
QY 1604 agtacacccttagaagactggaagagataaaaaagatagaacagctctctgagactcgg 1663
Db 3757 agtacacccttagaagactggaagagataaaaaagatagaacagctctctgagactcgg 3816
QY 1664 aggggacctgtctatctatctatctatctatctatctatctatctatctatctatctat 1723
Db 3817 aggggacctgtctatctatctatctatctatctatctatctatctatctatctatctat 3876
QY 1724 atgtccaaacagatatacgttcttaaccaccttagaagaataacacctgacagctgtgtca 1783
Db 3877 atgtccaaacagatatacgttcttaaccaccttagaagaataacacctgacagctgtgtca 3936
QY 1784 gaacatcccttctatctatctatctatctatctatctatctatctatctatctatctat 1843
Db 3937 gaacatcccttctatctatctatctatctatctatctatctatctatctatctatctat 3996
QY 1844 gtgtccctctgacagctggttaagtcaatcccttccagaaatacctccaggaagagactaa 1903
Db 3997 gtgtccctctgacagctggttaagtcaatcccttccagaaatacctccaggaagagactaa 4056
QY 1904 gggggaagccaaacccagcgcctactggaagtgagacttcaatgagtaaaagccgagtaaat 1963
Db 4057 gggggaagccaaacccagcgcctactggaagtgagacttcaatgagtaaaagccgagtaaat 4116
QY 1964 acggaacaaatactatctatctatctatctatctatctatctatctatctatctatctat 2023
Db 4117 acggaacaaatactatctatctatctatctatctatctatctatctatctatctatctat 4176
QY 2024 ctactaagaagaagacttcaacccgtgtgtctaaagaatactctgaggaataattttccaa 2083
Db 4177 ctactaagaagaagacttcaacccgtgtgtctaaagaatactctgaggaataattttccaa 4236
QY 2084 gatttggaaatacctaaagtaataatggttcagaaatgttccagcttctgttccagagtaa 2143
Db 4237 gatttggaaatacctaaagtaataatggttcagaaatgttccagcttctgttccagagtaa 4296
QY 2144 gtaagggactggtccaaagatactgggagtgatgtgaactgcatgtgcatatacagaagccc 2203
Db 4297 gtaagggactggtccaaagatactgggagtgatgtgaactgcatgtgcatatacagaagccc 4356
QY 2204 aaagctcagaagcagtagagaagatgaatagaacccataaagaagacccttaaccaatgtga 2263
Db 4357 aaagctcagaagcagtagagaagatgaatagaacccataaagaagacccttaaccaatgtga 4416
QY 2264 ccaagagactggtcattaaatgaatgattgattgattgattgattgattgattgattgattg 2323
Db 4417 ccaagagactggtcattaaatgaatgattgattgattgattgattgattgattgattgattg 4476
QY 2324 ggaacacccctgacagcttgggctgacccctataatgtctcctcctgttcttaagggtga 2383
Db 4477 ggaacacccctgacagcttgggctgacccctataatgtctcctcctgttcttaagggtga 4536
QY 2384 cgttggcaagaatgtccttgcacatagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2443
Db 4537 cgttggcaagaatgtccttgcacatagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4596
QY 2444 ctgagctcaagcgcctcagatggtgtgagcagcagcgtggtgaagcagctccggagagcct 2503
Db 4597 ctgagctcaagcgcctcagatggtgtgagcagcagcgtggtgaagcagctccggagagcct 4656
QY 2504 acccagagagagagctgacagcttccacatgccttccaagtgtgagattagttatgtta 2563

```

```

Db 4657 actcagagagagactgtcgaatccacatcgtcttccaaagttgagatcagttatgtta 4716
QY 2564 gacgccaccgtgcaagaacctcgcagactcgtgtgaaggaggaacctatctcgttacttga 2623
Db 4717 gacgccaccgtgcaagaacctcgcagactcgtgtgaaggaggaacctatctcgttacttga 4776
QY 2624 ccacaccaagcgtgtgaaagtcgaaggaatcccc 2658
Db 4777 ccacaccaagcgtgtgaaagtcgaaggaatcccc 4811

RESULT 5
AAT74883
ID AAT74883 standard; cDNA: 7393 BP.
XX
AC AAT74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key
FH CDS Location/Qualifiers
FH 598..2172
FH /*tag= a
FH 598..2169
FH /*tag= b
FH /*tag= c
FH /*tag= d
FH /*tag= e
FH /*tag= f
FH /*tag= g
FH /*tag= h
FH /*tag= i
FH /*tag= j
FH /*tag= k
FH /*tag= l
FH /*tag= m
FH /*tag= n
FH /*tag= o
FH /*tag= p
FH /*tag= q
FH /*tag= r
FH /*tag= s
FH /*tag= t
FH /*tag= u
FH /*tag= v
FH /*tag= w
FH /*tag= x
FH /*tag= y
FH /*tag= z
FH /*tag= aa
FH /*tag= ab
FH /*tag= ac
FH /*tag= ad
FH /*tag= ae
FH /*tag= af
FH /*tag= ag
FH /*tag= ah
FH /*tag= ai
FH /*tag= aj
FH /*tag= ak
FH /*tag= al
FH /*tag= am
FH /*tag= an
FH /*tag= ao
FH /*tag= ap
FH /*tag= aq
FH /*tag= ar
FH /*tag= as
FH /*tag= at
FH /*tag= au
FH /*tag= av
FH /*tag= aw
FH /*tag= ax
FH /*tag= ay
FH /*tag= az
FH /*tag= ba
FH /*tag= bb
FH /*tag= bc
FH /*tag= bd
FH /*tag= be
FH /*tag= bf
FH /*tag= bg
FH /*tag= bh
FH /*tag= bi
FH /*tag= bj
FH /*tag= bk
FH /*tag= bl
FH /*tag= bm
FH /*tag= bn
FH /*tag= bo
FH /*tag= bp
FH /*tag= bq
FH /*tag= br
FH /*tag= bs
FH /*tag= bt
FH /*tag= bu
FH /*tag= bv
FH /*tag= bw
FH /*tag= bx
FH /*tag= by
FH /*tag= bz
FH /*tag= ca
FH /*tag= cb
FH /*tag= cc
FH /*tag= cd
FH /*tag= ce
FH /*tag= cf
FH /*tag= cg
FH /*tag= ch
FH /*tag= ci
FH /*tag= cj
FH /*tag= ck
FH /*tag= cl
FH /*tag= cm
FH /*tag= cn
FH /*tag= co
FH /*tag= cp
FH /*tag= cq
FH /*tag= cr
FH /*tag= cs
FH /*tag= ct
FH /*tag= cu
FH /*tag= cv
FH /*tag= cw
FH /*tag= cx
FH /*tag= cy
FH /*tag= cz
FH /*tag= da
FH /*tag= db
FH /*tag= dc
FH /*tag= dd
FH /*tag= de
FH /*tag= df
FH /*tag= dg
FH /*tag= dh
FH /*tag= di
FH /*tag= dj
FH /*tag= dk
FH /*tag= dl
FH /*tag= dm
FH /*tag= dn
FH /*tag= do
FH /*tag= dp
FH /*tag= dq
FH /*tag= dr
FH /*tag= ds
FH /*tag= dt
FH /*tag= du
FH /*tag= dv
FH /*tag= dw
FH /*tag= dx
FH /*tag= dy
FH /*tag= dz
FH /*tag= ea
FH /*tag= eb
FH /*tag= ec
FH /*tag= ed
FH /*tag= ee
FH /*tag= ef
FH /*tag= eg
FH /*tag= eh
FH /*tag= ei
FH /*tag= ej
FH /*tag= ek
FH /*tag= el
FH /*tag= em
FH /*tag= en
FH /*tag= eo
FH /*tag= ep
FH /*tag= eq
FH /*tag= er
FH /*tag= es
FH /*tag= et
FH /*tag= eu
FH /*tag= ev
FH /*tag= ew
FH /*tag= ex
FH /*tag= ey
FH /*tag= ez
FH /*tag= fa
FH /*tag= fb
FH /*tag= fc
FH /*tag= fd
FH /*tag= fe
FH /*tag= ff
FH /*tag= fg
FH /*tag= fh
FH /*tag= fi
FH /*tag= fj
FH /*tag= fk
FH /*tag= fl
FH /*tag= fm
FH /*tag= fn
FH /*tag= fo
FH /*tag= fp
FH /*tag= fq
FH /*tag= fr
FH /*tag= fs
FH /*tag= ft
FH /*tag= fu
FH /*tag= fv
FH /*tag= fw
FH /*tag= fx
FH /*tag= fy
FH /*tag= fz
FH /*tag= ga
FH /*tag= gb
FH /*tag= gc
FH /*tag= gd
FH /*tag= ge
FH /*tag= gf
FH /*tag= gg
FH /*tag= gh
FH /*tag= gi
FH /*tag= gj
FH /*tag= gk
FH /*tag= gl
FH /*tag= gm
FH /*tag= gn
FH /*tag= go
FH /*tag= gp
FH /*tag= gq
FH /*tag= gr
FH /*tag= gs
FH /*tag= gt
FH /*tag= gu
FH /*tag= gv
FH /*tag= gw
FH /*tag= gx
FH /*tag= gy
FH /*tag= gz
FH /*tag= ha
FH /*tag= hb
FH /*tag= hc
FH /*tag= hd
FH /*tag= he
FH /*tag= hf
FH /*tag= hg
FH /*tag= hh
FH /*tag= hi
FH /*tag= hj
FH /*tag= hk
FH /*tag= hl
FH /*tag= hm
FH /*tag= hn
FH /*tag= ho
FH /*tag= hp
FH /*tag= hq
FH /*tag= hr
FH /*tag= hs
FH /*tag= ht
FH /*tag= hu
FH /*tag= hv
FH /*tag= hw
FH /*tag= hx
FH /*tag= hy
FH /*tag= hz
FH /*tag= ia
FH /*tag= ib
FH /*tag= ic
FH /*tag= id
FH /*tag= ie
FH /*tag= if
FH /*tag= ig
FH /*tag= ih
FH /*tag= ii
FH /*tag= ij
FH /*tag= ik
FH /*tag= il
FH /*tag= im
FH /*tag= in
FH /*tag= io
FH /*tag= ip
FH /*tag= iq
FH /*tag= ir
FH /*tag= is
FH /*tag= it
FH /*tag= iu
FH /*tag= iv
FH /*tag= iw
FH /*tag= ix
FH /*tag= iy
FH /*tag= iz
FH /*tag= ja
FH /*tag= jb
FH /*tag= jc
FH /*tag= jd
FH /*tag= je
FH /*tag= jf
FH /*tag= jg
FH /*tag= jh
FH /*tag= ji
FH /*tag= jj
FH /*tag= jk
FH /*tag= jl
FH /*tag= jm
FH /*tag= jn
FH /*tag= jo
FH /*tag= jp
FH /*tag= jq
FH /*tag= jr
FH /*tag= js
FH /*tag= jt
FH /*tag= ju
FH /*tag= jv
FH /*tag= jw
FH /*tag= jx
FH /*tag= jy
FH /*tag= jz
FH /*tag= ka
FH /*tag= kb
FH /*tag= kc
FH /*tag= kd
FH /*tag= ke
FH /*tag= kf
FH /*tag= kg
FH /*tag= kh
FH /*tag= ki
FH /*tag= kj
FH /*tag= kk
FH /*tag= kl
FH /*tag= km
FH /*tag= kn
FH /*tag= ko
FH /*tag= kp
FH /*tag= kq
FH /*tag= kr
FH /*tag= ks
FH /*tag= kt
FH /*tag= ku
FH /*tag= kv
FH /*tag= kw
FH /*tag= kx
FH /*tag= ky
FH /*tag= kz
FH /*tag= la
FH /*tag= lb
FH /*tag= lc
FH /*tag= ld
FH /*tag= le
FH /*tag= lf
FH /*tag= lg
FH /*tag= lh
FH /*tag= li
FH /*tag= lj
FH /*tag= lk
FH /*tag= ll
FH /*tag= lm
FH /*tag= ln
FH /*tag= lo
FH /*tag= lp
FH /*tag= lq
FH /*tag= lr
FH /*tag= ls
FH /*tag= lt
FH /*tag= lu
FH /*tag= lv
FH /*tag= lw
FH /*tag= lx
FH /*tag= ly
FH /*tag= lz
FH /*tag= ma
FH /*tag= mb
FH /*tag= mc
FH /*tag= md
FH /*tag= me
FH /*tag= mf
FH /*tag= mg
FH /*tag= mh
FH /*tag= mi
FH /*tag= mj
FH /*tag= mk
FH /*tag= ml
FH /*tag= mn
FH /*tag= mo
FH /*tag= mp
FH /*tag= mq
FH /*tag= mr
FH /*tag= ms
FH /*tag= mt
FH /*tag= mu
FH /*tag= mv
FH /*tag= mw
FH /*tag= mx
FH /*tag= my
FH /*tag= mz
FH /*tag= na
FH /*tag= nb
FH /*tag= nc
FH /*tag= nd
FH /*tag= ne
FH /*tag= nf
FH /*tag= ng
FH /*tag= nh
FH /*tag= ni
FH /*tag= nj
FH /*tag= nk
FH /*tag= nl
FH /*tag= nm
FH /*tag= no
FH /*tag= np
FH /*tag= nq
FH /*tag= nr
FH /*tag= ns
FH /*tag= nt
FH /*tag= nu
FH /*tag= nv
FH /*tag= nw
FH /*tag= nx
FH /*tag= ny
FH /*tag= nz
FH /*tag= oa
FH /*tag= ob
FH /*tag= oc
FH /*tag= od
FH /*tag= oe
FH /*tag= of
FH /*tag= og
FH /*tag= oh
FH /*tag= oi
FH /*tag= oj
FH /*tag= ok
FH /*tag= ol
FH /*tag= om
FH /*tag= on
FH /*tag= oo
FH /*tag= op
FH /*tag= oq
FH /*tag= or
FH /*tag= os
FH /*tag= ot
FH /*tag= ou
FH /*tag= ov
FH /*tag= ow
FH /*tag= ox
FH /*tag= oy
FH /*tag= oz
FH /*tag= pa
FH /*tag= pb
FH /*tag= pc
FH /*tag= pd
FH /*tag= pe
FH /*tag= pf
FH /*tag= pg
FH /*tag= ph
FH /*tag= pi
FH /*tag= pj
FH /*tag= pk
FH /*tag= pl
FH /*tag= pm
FH /*tag= pn
FH /*tag= po
FH /*tag= pp
FH /*tag= pq
FH /*tag= pr
FH /*tag= ps
FH /*tag= pt
FH /*tag= pu
FH /*tag= pv
FH /*tag= pw
FH /*tag= px
FH /*tag= py
FH /*tag= pz
FH /*tag= qa
FH /*tag= qb
FH /*tag= qc
FH /*tag= qd
FH /*tag= qe
FH /*tag= qf
FH /*tag= qg
FH /*tag= qh
FH /*tag= qi
FH /*tag= qj
FH /*tag= qk
FH /*tag= ql
FH /*tag= qm
FH /*tag= qn
FH /*tag= qo
FH /*tag= qp
FH /*tag= qq
FH /*tag= qr
FH /*tag= qs
FH /*tag= qt
FH /*tag= qu
FH /*tag= qv
FH /*tag= qw
FH /*tag= qx
FH /*tag= qy
FH /*tag= qz
FH /*tag= ra
FH /*tag= rb
FH /*tag= rc
FH /*tag= rd
FH /*tag= re
FH /*tag= rf
FH /*tag= rg
FH /*tag= rh
FH /*tag= ri
FH /*tag= rj
FH /*tag= rk
FH /*tag= rl
FH /*tag= rm
FH /*tag= rn
FH /*tag= ro
FH /*tag= rp
FH /*tag= rq
FH /*tag= rr
FH /*tag= rs
FH /*tag= rt
FH /*tag= ru
FH /*tag= rv
FH /*tag= rw
FH /*tag= rx
FH /*tag= ry
FH /*tag= rz
FH /*tag= sa
FH /*tag= sb
FH /*tag= sc
FH /*tag= sd
FH /*tag= se
FH /*tag= sf
FH /*tag= sg
FH /*tag= sh
FH /*tag= si
FH /*tag= sj
FH /*tag= sk
FH /*tag= sl
FH /*tag= sm
FH /*tag= sn
FH /*tag= so
FH /*tag= sp
FH /*tag= sq
FH /*tag= sr
FH /*tag= ss
FH /*tag= st
FH /*tag= su
FH /*tag= sv
FH /*tag= sw
FH /*tag= sx
FH /*tag= sy
FH /*tag= sz
FH /*tag= ta
FH /*tag= tb
FH /*tag= tc
FH /*tag= td
FH /*tag= te
FH /*tag= tf
FH /*tag= tg
FH /*tag= th
FH /*tag= ti
FH /*tag= tj
FH /*tag= tk
FH /*tag= tl
FH /*tag= tm
FH /*tag= tn
FH /*tag= to
FH /*tag= tp
FH /*tag= tq
FH /*tag= tr
FH /*tag= ts
FH /*tag= tt
FH /*tag= tu
FH /*tag= tv
FH /*tag= tw
FH /*tag= tx
FH /*tag= ty
FH /*tag= tz
FH /*tag= ua
FH /*tag= ub
FH /*tag= uc
FH /*tag= ud
FH /*tag= ue
FH /*tag= uf
FH /*tag= ug
FH /*tag= uh
FH /*tag= ui
FH /*tag= uj
FH /*tag= uk
FH /*tag= ul
FH /*tag= um
FH /*tag= un
FH /*tag= uo
FH /*tag= up
FH /*tag= uq
FH /*tag= ur
FH /*tag= us
FH /*tag= ut
FH /*tag= uv
FH /*tag= uw
FH /*tag= ux
FH /*tag= uy
FH /*tag= uz
FH /*tag= va
FH /*tag= vb
FH /*tag= vc
FH /*tag= vd
FH /*tag= ve
FH /*tag= vf
FH /*tag= vg
FH /*tag= vh
FH /*tag= vi
FH /*tag= vj
FH /*tag= vk
FH /*tag= vl
FH /*tag= vm
FH /*tag= vn
FH /*tag= vo
FH /*tag= vp
FH /*tag= vq
FH /*tag= vr
FH /*tag= vs
FH /*tag= vt
FH /*tag= vu
FH /*tag= vv
FH /*tag= vw
FH /*tag= vx
FH /*tag= vy
FH /*tag= vz
FH /*tag= wa
FH /*tag= wb
FH /*tag= wc
FH /*tag= wd
FH /*tag= we
FH /*tag= wf
FH /*tag= wg
FH /*tag= wh
FH /*tag= wi
FH /*tag= wj
FH /*tag= wk
FH /*tag= wl
FH /*tag= wm
FH /*tag= wn
FH /*tag= wo
FH /*tag= wp
FH /*tag= wq
FH /*tag= wr
FH /*tag= ws
FH /*tag= wt
FH /*tag= wu
FH /*tag= wv
FH /*tag= ww
FH /*tag= wx
FH /*tag= wy
FH /*tag= wz
FH /*tag= xa
FH /*tag= xb
FH /*tag= xc
FH /*tag= xd
FH /*tag= xe
FH /*tag= xf
FH /*tag= xg
FH /*tag= xh
FH /*tag= xi
FH /*tag= xj
FH /*tag= xk
FH /*tag= xl
FH /*tag= xm
FH /*tag= xn
FH /*tag= xo
FH /*tag= xp
FH /*tag= xq
FH /*tag= xr
FH /*tag= xs
FH /*tag= xt
FH /*tag= xu
FH /*tag= xv
FH /*tag= xw
FH /*tag= xx
FH /*tag= xy
FH /*tag= xz
FH /*tag= ya
FH /*tag= yb
FH /*tag= yc
FH /*tag= yd
FH /*tag= ye
FH /*tag= yf
FH /*tag= yg
FH /*tag= yh
FH /*tag= yi
FH /*tag= yj
FH /*tag= yk
FH /*tag= yl
FH /*tag= ym
FH /*tag= yn
FH /*tag= yo
FH /*tag= yp
FH /*tag= yq
FH /*tag= yr
FH /*tag= ys
FH /*tag= yt
FH /*tag= yu
FH /*tag= yv
FH /*tag= yw
FH /*tag= yx
FH /*tag= yy
FH /*tag= yz
FH /*tag= za
FH /*tag= zb
FH /*tag= zc
FH /*tag= zd
FH /*tag= ze
FH /*tag= zf
FH /*tag= zg
FH /*tag= zh
FH /*tag= zi
FH /*tag= zj
FH /*tag= zk
FH /*tag= zl
FH /*tag= zm
FH /*tag= zn
FH /*tag= zo
FH /*tag= zp
FH /*tag= zq
FH /*tag= zr
FH /*tag= zs
FH /*tag= zt
FH /*tag= zu
FH /*tag= zv
FH /*tag= zw
FH /*tag= zx
FH /*tag= zy
FH /*tag= zz

```

Claim 16; Fig 2; 128bp; English.

XX This cDNA sequence represents a defective purified swine retrovirus  
 CC found in PK-15 cells containing the putative coding regions for viral  
 CC GAG, POL and ENV proteins. There are a few in frame stop codons and  
 CC apparent frame shifts in the given coding sequence which alter features  
 CC of the translation. This sequence and PCR fragments generated from the  
 CC sequence (see A474812-T74882) could be used to screen organs for the  
 CC presence of porcine retroviruses prior to xenotransplantation.  
 CC Transplantation can increase the likelihood of retroviral activation if  
 CC intact and infectious proviruses are present. The porcine retroviral  
 CC sequence can be used to generate probes to determine the level (e.g.  
 CC copy number) of intact (i.e. potentially replicating) porcine proviruses  
 CC sequences in a strain of xenograft transplantation donors. It can be  
 CC used to detect mutations, genetic lesions or viral recombinants and to  
 CC determine the histological localisation of activated retroviruses. Using  
 CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear  
 CC cells, infectivity titration and susceptibility testing can be  
 CC performed. Ultimately animal donors without intact porcine retroviral  
 CC sequences or a lower copy number of viral elements could be selected.  
 XX  
 SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

Query Match 32.9%; Score 1091; DB 18; Length 7393;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1291; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1364 aaaggggtgtcttaacctagcagggaggaataaagaacaagaagaaatttaagcc 1423  
 DB 3517 aaaggggtgtcttaacctagcagggaggaataaagaacaagaagaaatttaagcc 3576  
 QY 1424 tattagaagccttattgccaagaagcgtagctattatatactgtccctggaaataga 1483  
 DB 3577 tattagaagccttattgccaagaagcgtagctattatatactgtccctggaaataga 3636  
 QY 1484 aagcacaagatcatalctagaggaaccagatgctgacgggtgccaagcagcg 1543  
 DB 3637 aagcacaagatcatalctagaggaaccagatgctgacgggtgccaagcagcg 3696  
 QY 1544 cccaggtgttaacctctgcttaataagaacgcgccaaagccccaagcccgagcag 1603  
 DB 3697 cccaggtgttaacctctgcttaataagaacgcgccaaagccccaagcccgagcag 3756  
 QY 1604 agtacacctagaagactggaagagaataaaagaatagaccggttcttgagactcgg 1663  
 DB 3757 agtacacctagaagactggaagagaataaaagaatagaccggttcttgagactcgg 3816  
 QY 1664 agggagacctgtatactcatalatggaaggaatacctgcgccacaagaaggggttaaat 1723  
 DB 3817 agggagacctgtatactcatalatggaaggaatacctgcgccacaagaaggggttaaat 3876  
 QY 1724 atgtccacaagatatacgtcttaaccacactagaactaaacacctgcagctgtgta 1783  
 DB 3877 atgtccacaagatatacgtcttaaccacactagaactaaacacctgcagctgtgta 3936  
 QY 1784 gaacatcccttatacatgttcttgaggtaccagagtggtgctgactcgtgtgtaaat 1843  
 DB 3937 gaacatcccttatacatgttcttgaggtaccagagtggtgctgactcgtgtgtaaat 3996  
 QY 1844 gtgtgacctgacagctgttaatagtatactcctccagaataactccagaagaagagctaa 1903  
 DB 3997 gtgtgacctgacagctgttaatagtatactcctccagaataactccagaagaagagctaa 4056  
 QY 1904 gggaagacacacacagcgcgtcaactgggaagtgaactcactgagtaaaagccgctaaat 1963  
 DB 4057 gggaagacacacacagcgcgtcaactgggaagtgaactcactgagtaaaagccgctaaat 4116  
 QY 1964 acggaacaataatctatgtttgtttgtagacaccttttagagatggtagagctatc 2023  
 DB 4117 acggaacaataatctatgtttgtttgtagacaccttttagagatggtagagctatc 4176  
 QY 2024 ctactaagaagagactcaacgltgtgtgtaagaataactcgtgggaatattccaa 2083  
 DB 2083 ctactaagaagagactcaacgltgtgtgtaagaataactcgtgggaatattccaa 2083

DB 4177 ctactaagaagagactcaacgltgtgtgtaagaataactcgtgggaatattccaa 4236  
 QY 2084 gatttgaataacctaaagttataggtcagacaatgttccagcttctgtcccaagtaa 2143  
 DB 4237 gatttgaataacctaaagttataggtcagacaatgttccagcttctgtcccaagtaa 4296  
 QY 2144 gtccagagcgtgccaagaatattgggattgtattggaaacgtcaattgtgataagaccc 2203  
 DB 4297 gtccagagcgtgccaagaatattgggattgtattggaaacgtcaattgtgataagaccc 4356  
 QY 2204 aaagctagaagacggtatgaagaatgatgatgaacccattaaagagacccttaacaattga 2263  
 DB 4357 aaagctagaagacggtatgaagaatgatgatgaacccattaaagagacccttaacaattga 4416  
 QY 2264 ccacagagactggaatgaatgatgatgtggtctcctcctgtgctttttaggttga 2323  
 DB 4417 ccacagagactggaatgaatgatgatgtggtctcctcctgtgctttttaggttga 4476  
 QY 2324 ggaacaccctgtgacaggtttgggtgtgacccctatgaattgtccttaaggggacccccc 2383  
 DB 4477 ggaacaccctgtgacaggtttgggtgtgacccctatgaattgtccttaaggggacccccc 4536  
 QY 2384 cgttggcagaatgtgctcttgacatagtgatgtgtgctgtcttccagccctgttct 2443  
 DB 4537 cgttggcagaatgtgctcttgacatagtgatgtgtgctgtcttccagccctgttct 4596  
 QY 2444 cttaggtcaaagcgtctcgtgagtggtgtagcagcagcgttggaaacagctcccggaagcct 2503  
 DB 4597 cttaggtcaaagcgtctcgtgagtggtgtagcagcagcgttggaaacagctcccggaagcct 4656  
 QY 2504 acccagaagagagactgtgcaagttccacatcgtcttccaagtgtgagatltcaatcgtatga 2563  
 DB 4657 acccagaagagagactgtgcaagttccacatcgtcttccaagtgtgagatltcaatcgtatga 4716  
 QY 2564 gacgcacacgttgcaggaacacctgcagactcgtgtggaagagacttatctcgtactttga 2623  
 DB 4717 gacgcacacgttgcaggaacacctgcagactcgtgtggaagagacttatctcgtactttga 4776  
 QY 2624 ccacaccacagcgtctgtaaaagtcgaaggaatcccc 2658  
 DB 4777 ccacaccacagcgtctgtaaaagtcgaaggaatcccc 4811

RESULT 6  
 AAC67020  
 ID AAC67020 standard; DNA; 4402 BP.  
 XX  
 AC AAC67020;  
 XX  
 XX 27-MAR-2001 (first entry)  
 DT  
 XX PERV env protein coding sequence SEQ ID NO: 20.  
 DE  
 XX Xenotransplantation; infectious agent; vaccine; ds.  
 KW  
 XX Porcine endogenous retrovirus.  
 OS  
 XX WO200071726-A1.  
 PN  
 XX 30-NOV-2000.  
 PD  
 XX 24-MAY-2000; 2000WO-US14296.  
 PF  
 XX 24-MAY-1999; 99US-0135631.  
 PR  
 XX (MAYO-) MAYO MEDICAL VENTURES.  
 PA  
 XX Federspiel MJ;  
 PI  
 XX WPI; 2001-032041/04.  
 DR  
 XX Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
XX  
PS Claim 16, page 105-106; 144pp; English.

XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.

XX Sequence 4402 BP; 1259 A; 1085 C; 1111 G; 947 T; 0 other;

Query Match 25.5%; Score 848; DB 22; Length 4402;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 2198; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 31 tctctcgtgatttaacccactagccaacacttttgcttgcgtgaatgagatcca 90  
DB 1043 tctcgtcgtgatttaacccactagccaacacttttgcttgcgtgaatgagatcca 1102  
QY 91 ggtacgggaagaacccgggagctcactgacccgactgcccgaaggttccaagaatcc 150  
DB 1103 ggtacgggaagaacccgggagctcactgacccgactgcccgaaggttccaagaatcc 1162  
QY 151 ccgacacactcttgacgaagccctacacagagaccctggtcccaacttcagatccacaccc 210  
DB 1163 ccgacacactcttgacgaagccctacacagagaccctggtcccaacttcagatccacaccc 1222  
QY 211 caagtgacccctctccagtcacgttggatgacctctcttgcgggagccacacacagac 270  
DB 1223 caagtgacccctctccagtcacgttggatgacctctcttgcgggagccacacacagac 1282  
QY 271 tgcctaaaggtacgaaggaactactctgtaattgctactactagctacagactc 330  
DB 1283 tgcctaaaggtacgaaggaactactctgtaattgctactactagctacagactc 1342  
QY 331 gctaaagaagccacagatttcagaggaagagtaacatacttgggtacagtttcgagac 390  
DB 1343 gctaaagaagccacagatttcagaggaagagtaacatacttgggtacagtttcgagac 1402  
QY 391 gggcagcagatgctcagcagagcagcagaaagaacttgaatcagataccggcccaacc 450  
DB 1403 gggcagcagatgctcagcagagcagcagaaagaacttgaatcagataccggcccaacc 1462  
QY 451 acacccaaacaagtgagagagtttttgggacagctggaatttcgacagactgtgacccg 510  
DB 1463 acacccaaacaagtgagagagtttttgggacagctggaatttcgacagactgtgacccg 1522  
QY 511 gggtttgcagacttagcagcccaactacccgtaacaaagaaggggagattcc 570  
DB 1523 gggtttgcagacttagcagcccaactacccgtaacaaagaaggggagattcc 1582  
QY 571 tgggctctcagcagcagagacatttgatgtatacaaaagccctggcagacgacac 630  
DB 1583 tgggctctcagcagcagagacatttgatgtatacaaaagccctggcagacgacac 1642  
QY 631 gctctggccctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 690  
DB 1643 gctctggccctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1702  
QY 691 gtagccgagagattttaacccaaacccatagagacatgtagagagacactgttgcactact 750  
DB 1703 gtagccgagagattttaacccaaacccatagagacatgtagagagacactgttgcactact 1762  
QY 751 tcaagaagacttgatctctgttagccagtggttggccgtatgctggaaggtatccgagct 810  
DB 1763 tcaagaagacttgatctctgttagccagtggttggccgtatgctggaaggtatccgagct 1822  
QY 811 gtggcctactgtcgaagagacgtgacaatctgacttgggacagaaatacaactgtata 870  
DB 1823 gtggcctactgtcgaagagacgtgacaatctgacttgggacagaaatacaactgtata 1882

QY 871 gcccccacatgcttgagagacatcgttcgcagccccacagaccgatgataccaacgccc 930  
DB 1883 gcccccacatgcttgagagacatcgttcgcagccccacagaccgatgataccaacgccc 1942  
QY 931 cgcataccacatcaaaagcctgtcttccagagagaggtcacttctgcctccacagccc 990  
DB 1943 cgcataccacatcaaaagcctgtcttccagagagaggtcacttctgcctccacagccc 2002  
QY 991 gcttcaacccctgcacactctctgccttgaagaagactgtatgaacagctgactatgtgc 1050  
DB 2003 gcttcaacccctgcacactctctgccttgaagaagactgtatgaacagctgactatgtgc 2062  
QY 1051 catcaactatgtatgagagagacttgggtccgcaaggaacttcaagacataccgtgact 1110  
DB 2063 catcaactatgtatgagagagacttgggtccgcaaggaacttcaagacataccgtgact 2122  
QY 1111 ggaagaaggtactactgttcaacttgaagaagcagctatgttgggtgaaggtgaagaagt 1170  
DB 2123 ggaagaaggtactactgttcaacttgaagaagcagctatgttgggtgaaggtgaagaagt 2182  
QY 1171 gcttgggcgagctgttgaacgggacccgcaagatcttggccagagccttgcggaagga 1230  
DB 2183 gcttgggcgagctgttgaacgggacccgcaagatcttggccagagccttgcggaagga 2242  
QY 1231 acttaagcgcaaaagcctgagctatgtgcctcaagcaagcttgcggtgcggaaggg 1290  
DB 2243 acttaagcgcaaaagcctgagctatgtgcctcaagcaagcttgcggtgcggaaggg 2302  
QY 1291 aatccataaacttatacgcagacagagatgtccttgcagactgcacagctaacaggg 1350  
DB 2303 aatccataaacttatacgcagacagagatgtccttgcagactgcacagctaacaggg 2362  
QY 1351 gccatctataaacaaggggtgttgcctaccccaagggagggaaataagaacaagaag 1410  
DB 2363 gccatctataaacaaggggtgttgcctaccccaagggagggaaataagaacaagaag 2422  
QY 1411 gaattctaaagcctattgaaagccttaccatttggccaaaaggtcgtcattatacagct 1470  
DB 2423 gaattctaaagcctattgaaagccttaccatttggccaaaaggtcgtcattatacagct 2482  
QY 1471 ccttgaacatcagaagaagccaaagatctcatatctaaagaagaaacagatgctgacccggt 1530  
DB 2483 ccttgaacatcagaagaagccaaagatctcatatctaaagaagaaacagatgctgacccggt 2542  
QY 1531 gccaaagcagccagccagcgttgaaccttgcctataatagaacaagcccaagccca 1590  
DB 2543 gccaaagcagccagccagcgttgaaccttgcctataatagaacaagcccaagccca 2602  
QY 1591 gaaccagcagcagatcaaccccttaagaagacttgcagagatataaagaatagacagctc 1650  
DB 2603 gaaccagcagcagatcaaccccttaagaagacttgcagagatataaagaatagacagctc 2662  
QY 1651 tctgaagactcggagagagcctcgtctatacctcatacttgggaaggaatccttgcacacaa 1710  
DB 2663 tctgaagactcggagagagcctcgtctatacctcatacctgaagtgggaaggaatccttgcacacaa 2722  
QY 1711 gaaaggttagaatatgtccacacagatatacgttcaaccactagaactaacaacctg 1770  
DB 2723 gaaaggttagaatatgtccacacagatatacgttcaaccactagaactaacaacctg 2782  
QY 1771 cagcagcttgcataagaacatctccttatacagcttctcgaagcttaccagaggttgcgtacccg 1830  
DB 2783 cagcagcttgcataagaacatctccttatacagcttctcgaagcttaccagaggttgcgtacccg 2842  
QY 1831 gtgttcaaacatctgttgcctcgcagcagcttgaatgttaactccttccagaataactcca 1890  
DB 2843 gtgttcaaacatctgttgcctcgcagcagcttgaatgttaactccttccagaataactcca 2902  
QY 1891 ggaagaagactaaggggaagccacacagcgtctacgttgggaagtgtgacttcaactgaagta 1950  
DB 2903 ggaagaagactaaggggaagccacacagcgtctacgttgggaagtgtgacttcaactgaagta 2962



QY 1951 aagcgcgctaatacaggaacaataatctatgtgtttttagagacacttttcagatg 2010  
|||||  
Db 2963 aagccgctaatacaggaacaataatctatgtgtttttagagacacttttcagatg 3022  
QY 2011 gtagagcgctatctactaagaagaagacttcaaccgtggtgctaaagaatactatgag 2070  
|||||  
Db 3023 gtagagcgctatctactaagaagaagacttcaaccgtggtgctaaagaatactatgag 3082  
QY 2071 gaaattttccaagatttgaatacctaaggaaggtcagacaaggtccagcttc 2130  
|||||  
Db 3083 gaaattttccaagatttgaatacctaaggaaggtcagacaaggtccagcttc 3142  
QY 2131 gttgccagagtaagtcagggactggtccaagatatgtgggattgttgaaactgcatgt 2190  
|||||  
Db 3143 gttgccagagtaagtcagggactggtccaagatatgtgggattgttgaaactgcatgt 3202  
QY 2191 gatacagagcccaagctcagagacagtgtagagagatgaataagaacatlaagaagcc 2250  
|||||  
Db 3203 gatacagagcccaagctcagagacagtgtagagagatgaataagaacatlaagaagcc 3262  
QY 2251 ctac 2255  
|||||  
Db 3263 ctac 3267  
RESULT 7  
AAC67021  
ID AAC67021 standard; DNA; 6076 BP.  
XX AAC67021;  
AC AAC67021;  
DT 27-MAR-2001 (first entry)  
XX PERV env protein coding sequence SEQ ID NO: 21.  
DE  
XX Xenotransplantation; infectious agent; vaccine; ds.  
KW Porcine endogenous retrovirus.  
OS  
XX WO200071726-A1.  
PN  
XX 30-NOV-2000.  
PD  
XX 24-MAY-2000; 2000WO-US14296.  
PF  
XX 24-MAY-1999; 99US-0135631.  
PR  
XX (MAYO-) MAYO MEDICAL VENTURES.  
PA  
XX Federspiel MJ;  
PI  
XX WPI; 2001-032041/04.  
DR  
XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT Transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PR swine cells -  
PS  
XX Claim 16; Page 107-109; 144pp; English.  
CC The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX  
XX Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 24.0%; Score 797; DB 22; Length 6076;  
Best Local Similarity 98.7%; Pred. No. 3.2e-314;  
Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 31 tctgcctgaagattacacccccaactagccaacacttttgccttcgaatgagatcca 90  
|||||  
Db 3577 tctgcctgaagattacacccccaactagccaacacttttgccttcgaatgagatcca 3636  
QY 91 ggtacgggaagaacccgggcaagctacacttgacccagctgcccaagggtccaagactcc 150  
|||||  
Db 3637 ggtacgggaagaacccgggcaagctacacttgacccagctgcccaagggtccaagactcc 3696  
QY 151 ccgagccctctttgacgaagccctacacagagagcttgcccaacttcagatccaacact 210  
|||||  
Db 3697 ccgagccctctttgacgaagccctacacagagagcttgcccaacttcagatccaacact 3756  
QY 211 caggttgacccctccccaactgacgtgagatgacctgtcttcgcygagacccacaacagac 270  
|||||  
Db 3757 caggttgacccctccccaactgacgtgagatgacctgtcttcgcygagacccacaacagac 3816  
QY 271 tgccttaagaagtaacgaagccaactacgtcgtggaattgtctgaacctagctacagacct 330  
|||||  
Db 3817 tgccttaagaagtaacgaagccaactacgtcgtggaattgtctgaacctagctacagacct 3876  
QY 331 gctaaagaagccccaagatttcagagaagaaggttaacatacttggtgttaacatttcggygc 390  
|||||  
Db 3877 gctaaagaagccccaagatttcagagaagaaggttaacatacttggtgttaacatttcggygc 3936  
QY 391 ggcagcagatgctgacgagagcgacggaagaacactgtatccagataccggcccaacc 450  
|||||  
Db 3937 ggcagcagatgctgacgagagcgacggaagaacactgtatccagataccggcccaacc 3996  
QY 451 acaagccaacaagtgagagagcttttggggacagcttgagatttcgaagactgtgatcccg 510  
|||||  
Db 3997 acaagccaacaagtgagagagcttttggggacagcttgagatttcgaagactgtgatcccg 4056  
QY 511 gggcttgaccccttagacagcccaactctaccgcttaaccacaagaagaaggaggaattctcc 570  
|||||  
Db 4057 gggcttgaccccttagacagcccaactctaccgcttaaccacaagaagaaggaggaattctcc 4116  
QY 571 tgggtctctgaagcaccgaagagcatltgactgactatacaaaaggccctgctgagcgacct 630  
|||||  
Db 4117 tgggtctctgaagcaccgaagagcatltgactgactatacaaaaggccctgctgagcgacct 4176  
QY 631 gctctgacctccctgaagcgttaactaaaccttaaccttaagtgtgagtgaagcgtlaaggaa 680  
|||||  
Db 4177 gctctgacctccctgaagcgttaactaaaccttaaccttaagtgtgagtgaagcgtlaaggaa 4236  
QY 691 gtaagccgagaggttttaaccacaaccttagagacatgagagagacctgtgtctaccgt 750  
|||||  
Db 4237 gtaagccgagaggttttaaccacaaccttagagacatgagagagacctgtgtctaccgt 4296  
QY 751 tcaagaagactgtatccctgtagccagtggttgccgtagtctgtaagggtatccagact 810  
|||||  
Db 4297 tcaagaagactgtatccctgtagccagtggttgccgtagtctgtaagggtatccagact 4356  
QY 811 gtagcactgtgtcaagaagcgtgacaacttgactttggagacagatatgaactgtaata 870  
|||||  
Db 4357 gtagcactgtgtcaagaagcgtgacaacttgactttggagacagatatgaactgtaata 4416  
QY 871 gcccaccatgcatltgagaagacatcgtttggcgaaccccccagaccgatggtgaccacggc 930  
|||||  
Db 4417 gcccaccatgcatltgagaagacatcgtttggcgaaccccccagaccgatggtgaccacggc 4476  
QY 931 cgcatagcaccatacaaaacctgctctcaacagagaagggtacacttgcctccaccagcc 990  
|||||  
Db 4477 cgcatagcaccatacaaaacctgctctcaacagagaagggtacacttgcctccaccagcc 4536  
QY 991 gctctcaaccctgacactcttctgctgtaagaagactgtatgaacacagtgactatgattgc 1050  
|||||  
Db 4537 gctctcaaccctgacactcttctgctgtaagaagactgtatgaacacagtgactatgattgc 4596  
QY 1051 catcaactatgtattgagagagactgggtccggaagagacttaacagataccatccgtgact 1110  
|||||  
Db 4597 catcaactatgtattgagagagactgggtccggaagagacttaacagataccatccgtgact 4656  
QY 1111 gtagaagtgctaaactggttcaactgacggaagcagctatgtgtgtaaggtlaagagatg 1170

```

Db 4657 ggagaaagtgtaacctgctacacgacgaaagcagctatgtgtggaaggtaagagagatg 4716
Qy 1171 gctggggcgagctggtggaagcggcagcagctggtggccaaacagctgcccgaagga 1230
Db 4717 gctggggcgagctggtggaagcggcagcagctggtggccaaacagctgcccgaagga 4776
Qy 1231 actcagcgcaaaagcgtgagctcagcctcagcagcagcttgcgtgcgtgcgcaagg 1290
Db 4777 actcagcgcaaaagcgtgagctcagcctcagcagcagcttgcgtgcgtgcgcaagg 4836
Qy 1291 aaatccataaaccattatagcagcagcagctgcttgcgtgcgtgcagcagcagc 1350
Db 4837 aaatccataaaccattatagcagcagcagctgcttgcgtgcgtgcagcagcagc 4896
Qy 1351 gccatctataaagaaaggggtgctctacacacagcagcagcagcagcagcagcagc 1410
Db 4897 gccatctataaagaaaggggtgctctacacacagcagcagcagcagcagcagcagc 4956
Qy 1411 gaaattcctaagcctattagaagccttaccatttgcacaaagcgtagctatatacatt 1470
Db 4957 gaaattcctaagcctattagaagccttaccatttgcacaaagcgtagctatatacatt 5016
Qy 1471 ccttggaacatcagaagcccaagatcctatctagagggaaaccagatggtcagcgggt 1530
Db 5017 ccttggaacatcagaagcccaagatcctatcctagagggaaaccagatggtcagcgggt 5076
Qy 1531 gccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1590
Db 5077 gccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5136
Qy 1591 gaaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1650
Db 5137 gaaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5196
Qy 1651 tctgagatcccgaggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1710
Db 5197 tctgagatcccgaggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5256
Qy 1711 gaaaggttagaataatgtccaaacagatcagctgctaaacacagcagcagcagcagcagc 1770
Db 5257 gaaaggttagaataatgtccaaacagatcagctgctaaacacagcagcagcagcagcagc 5316
Qy 1771 cagcagttgtgcagaaacaccccttaccatgcttgcagcagcagcagcagcagcagcagc 1830
Db 5317 cagcagttgtgcagaaacaccccttaccatgcttgcagcagcagcagcagcagcagcagc 5376
Qy 1831 gtgttcaaacattgtgtgcctgtgcagcagcagcagcagcagcagcagcagcagcagc 1890
Db 5377 gtgttcaaacattgtgtgcctgtgcagcagcagcagcagcagcagcagcagcagcagc 5436
Qy 1891 ggaagagactcaagggagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1950
Db 5437 ggaagagactcaagggagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5496
Qy 1951 aagccggttaataacggaacaaatctatgtgttcttgtagaacaccttccagagtg 2010
Db 5497 aagccggttaataacggaacaaatctatgtgttcttgtagaacaccttccagagtg 5556
Qy 2011 gtagaggttactctactaaagaagagactcaacgctgtgtgtgtgtgtgtgtgtgtgtgt 2070
Db 5557 gtagaggttactctactaaagaagagactcaacgctgtgtgtgtgtgtgtgtgtgtgtgt 5616
Qy 2071 gaaatttccaaagatttgatagatcctaaagtaaggttaaggttaagcagcagcagcagc 2130
Db 5617 gaaatttccaaagatttgatagatcctaaagtaaggttaaggttaagcagcagcagcagc 5676
Qy 2131 gttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2190
Db 5677 gttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5736
Qy 2191 gcatcagaccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2250

```

```

Db 5737 gcatcagaccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5796
Qy 2251 cttac 2255
Db 5797 cttac 5801

RESULT 8
AAC67019
ID AAC67019 standard; DNA; 7362 BP.
XX
AC AAC67019;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 19.
XX
KW Xenotransplantation; infectious agent; vaccine; ds.
XX
XX Porcine endogenous retrovirus.
OS
PN MO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000MO-US14296.
XX
PR 24-MAY-1999; 99US-0135631.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
PT Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
PS Claim 16; Page 101-104; 144pp; English.
XX
CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
SQ Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

Query Match 24.0%; Score 797; DB 22; Length 7362;
Best local Similarity 98.7%; Pred. No. 3,2e-314;
Matches 2197; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

```
D 3882 tgccttagaaggcaggaagcactactctgctgaattgtctgcactagctacagagcctct 3941
Q 331 gctaaagaagccagagattgacagagagaggttaacactcttggtgtacagcttggtg 390
D 3942 gctaaagaagccagagattgacagagagaggttaacactcttggtgtacagcttggtg 4001
Q 391 gggcagcagatgctgacgagagcagcagaagaataactgtatgcagatacgcggcccaac 450
D 4002 gggcagcagatgctgacgagagcagcagaagaataactgtatgcagatacgcggcccaac 4061
Q 451 aacagccaacaacagtgagagagatttttgggagacagctgagcttctgcagactgtgacccg 510
D 4062 aacagccaacaacagtgagagagatttttgggagacagctgagcttctgcagactgtgacccg 4121
Q 511 gggattgacagcttaacagcccaactctacgcgttaacaaagaagaaggagattctcc 570
D 4122 gggattgacagcttaacagcccaactctacgcgttaacaaagaagaaggagattctcc 4181
Q 571 tgggctcctgagcagcaggaagcatttgaatgctatcaaaaaggccctgtgagcgacact 630
D 4182 tgggctcctgagcagcaggaagcatttgaatgctatcaaaaaggccctgtgagcgacact 4241
Q 631 gctctggccctccctccagcttaacacttaaccccttaactgtgagcgttaaggga 690
D 4242 gctctggccctccctccagcttaacacttaaccccttaactgtgagcgttaaggga 4301
Q 691 gtagcccgagagatttaacccaacccctagagacatgagagagacccgttgccactg 750
D 4302 gtagcccgagagatttaacccaacccctagagacatgagagagacccgttgccactg 4361
Q 751 tcaaaagaagcttgatccctgtgagcagtggttggccgctatgtctgaagctatccgact 810
D 4362 tcaaaagaagcttgatccctgtgagcagtggttggccgctatgtctgaagctatccgact 4421
Q 811 gtagccatactgtgacagcagctgacaaatgactttgggacaggaataataactgtata 870
D 4422 gtagccatactgtgacagcagctgacaaatgactttgggacaggaataataactgtata 4481
Q 871 gcccccactgcatctggaagaacatcgtctgcgagcccccagacagctgatacaaacgccc 930
D 4482 gcccccactgcatctggaagaacatcgtctgcgagcccccagacagctgatacaaacgccc 4541
Q 931 cgcatagacccactatcaaaagcctgctctcaagaagaaggtcacttgcctccacacagcc 990
D 4542 cgcatagacccactatcaaaagcctgctctcaagaagaaggtcacttgcctccacacagcc 4601
Q 991 gctcacaacccctgcaacccctctctgcgaagaagactgtagaacaagagaccatgattgc 1030
D 4602 gctcacaacccctgcaacccctctctgcgaagaagactgtagaacaagagaccatgattgc 4661
Q 1051 catcaactattgattgagagagactgggtctccgaagaagccttacagacataccgctgact 1110
D 4662 catcaactattgattgagagagactgggtctccgaagaagccttacagacataccgctgact 4721
Q 1111 gggagaagtgctaaactgttctactgagcggaaagcactatgtgtggaagtgaaagagatg 1170
D 4722 gggagaagtgctaaactgttctactgagcggaaagcactatgtgtggaagtgaaagagatg 4781
Q 1171 gctggggcgagcagtgtagagcggaacccgacagatctggcgcaaacacccgcccgaagga 1230
D 4782 gctggggcgagcagtgtagagcggaacccgacagatctggcgcaaacacccgcccgaagga 4841
Q 1231 acttcagcgcaaaaagcgtgagctcatgagcctcagcaaaccttgcgagctgcgcgaaggg 1290
D 4842 acttcagcgcaaaaagcgtgagctcatgagcctcagcaaaccttgcgagctgcgcgaaggg 4901
Q 1291 aaatccataaacatttatacggagacaggtatgcttgcgactgcacagctacacggg 1350
D 4902 aaatccataaacatttatacggagacaggtatgcttgcgactgcacagctacacggg 4961
Q 1351 gccatctataaacaagaagggttgcttaactcagcaggggggaaataaagaacaagaag 1410
D 4962 gccatctataaacaagaagggttgcttaactcagcaggggggaaataaagaacaagaag 5021

Q 1411 gaaattcaagccttatagaagccttatacttgcocaaaaagctagctattatacactgt 1470
D 5022 gaaattcaagccttatagaagccttatacttgcocaaaaagctagctattatacactgt 5081
Q 1471 cctgagacatcagaaagcccaagatctatctatgagaggaacccagatggtcagaccgggtc 1530
D 5082 cctgagacatcagaaagcccaagatctatctatgagaggaacccagatggtcagaccgggtc 5141
Q 1531 gccaaagcagcagcccaagcgtgttaacccctctgcctataatagaaacgcccgaagcccca 1590
D 5142 gccaaagcagcagcccaagcgtgttaacccctctgcctataatagaaacgcccgaagcccca 5201
Q 1591 gaaccccaacgacagctacacccctgaagaagcttgcagagagataaaaaagatagacagttc 1650
D 5202 gaaccccaacgacagctacacccctgaagaagcttgcagagagataaaaaagatagacagttc 5261
Q 1651 tctgagactccggaggggagccttgctatccctcatactgaggaaggaatccctgcggccacaaa 1710
D 5262 tctgagactccggaggggagccttgctatccctcatactgaggaaggaatccctgcggccacaaa 5321
Q 1711 gaaaggttagaataatgtccaaacagatatacgtctaaacccacactaggaagataaacactgt 1770
D 5322 gaaaggttagaataatgtccaaacagatatacgtctaaacccacactaggaagataaacactgt 5381
Q 1771 cagcagattgtgcagaaacatccctctatacagttcttgagagctaccaggaagtgtgactgc 1830
D 5382 cagcagattgtgcagaaacatccctctatacagttcttgagagctaccaggaagtgtgactgc 5441
Q 1831 gtagtcaacaatgtgtgcccctgcccagctgagctgtaagtcttaactctccaaataaccccca 1890
D 5442 gtagtcaacaatgtgtgcccctgcccagctgagctgtaagtcttaactctccaaataaccccca 5501
Q 1891 ggaagaagactaaagggaagccacccagcgcctcacttgggaagtggaacttcaactgagta 1950
D 5502 ggaagaagactaaagggaagccacccagcgcctcacttgggaagtggaacttcaactgagta 5561
Q 1951 aagccggcctaatacgaagaacaataatctatgtgttttgaacaccccttcaagatg 2010
D 5562 aagccggcctaatacgaagaacaataatctatgtgttttgaacaccccttcaagatg 5621
Q 2011 gtagagcttatacctataagaagaagagacttcaacgctgtgtgctaaagaataactggag 2070
D 5622 gtagagcttatacctataagaagaagagacttcaacgctgtgtgctaaagaataactggag 5681
Q 2071 gaaattttccaaagatttggaaatacctaagaagtaagtgctacagacaatggtccagcttc 2130
D 5682 gaaattttccaaagatttggaaatacctaagaagtaagtgctacagacaatggtccagcttc 5741
Q 2131 gttgccaggttaagtgacagagactgtgcgaagatatgttggtatgtggaactgactgt 2190
D 5742 gttgccaggttaagtgacagagactgtgcgaagatatgttggtatgtggaactgactgt 5801
Q 2191 gcatlacagaccccaagaagcttaagacaggttagaagaagtaaatgaacccattaaagagacc 2250
D 5802 gcatlacagaccccaagaagcttaagacaggttagaagaagtaaatgaacccattaaagagacc 5861
Q 2251 cttac 2255
D 5862 cttac 5866

RESULT 9
AAV82748
ID AAV82748 standard; DNA; 2462 BP.
XX
XX AAV82748;
AC
XX
XX 25-FEB-1999 (first entry)
DT
XX
XX pig endogenous retrovirus (PERV)-A envelope (env) gene region.
DE
XX pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
KW
```

KW probe: primer; detection: retrovirus; human tissue: xenotransplant;  
 KW primary porcine tissue; human cell line; porcine cell line; ss.  
 OS Pig endogenous retrovirus.

XX Key Location/Qualifiers  
 XX CDS 211..2193  
 XX /tag= a  
 XX /product= envelope-protein

PN W09853104-A2.

XX 26-NOV-1998.

XX 18-MAY-1998; 98MO-GB01428.

XX 16-MAY-1997; 97GB-0010154.

XX (MED1-) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI: 1999-045324/04.

DR P-PSDB; AAM85453.

PT Newly isolated nucleic acid probe capable of hybridizing to either  
 PT the PERV-A or PERV-B env gene - useful in the detection of  
 PT retroviruses, and their subtypes, in a sample of porcine/human  
 PT tissue

XX Claim 6; Page 20-21; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-A  
 CC envelope protein. PERV exists in two different subtypes, PERV-A and  
 CC PERV-B. The differences are reflected in sequence divergence in the  
 CC envelope genes. Probes and primers can be derived from the envelope  
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
 CC a method to detect retroviruses in a sample of porcine/human tissue,  
 CC particularly primary porcine tissue and human cell lines that have been  
 CC cultivated in the presence of a porcine cell line, or human tissue from  
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
 CC one of the PERV env genes can also be determined.

XX Sequence 2462 BP: 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match

23.7%; Score 786; DB 20; Length 2462;

Best Local Similarity 99.9%; Pred. No. 1,1e-309; Mismatches 1; Indels 0; Gaps 0;

Matches 836; Conservative 0;

OY 2459 tcgaatgggtgtagcagcagcgtggaagcagctccggagagcctactcagagagagact 2518  
 DB 1 ccgagtggtgtagcagcagcgtggaagcagctccggagagcctactcagagagagact 60  
 OY 2519 tggcaagttccacatcgttcccaagttgagagttcagttatgtagagcagccgtgag 2578  
 DB 61 tggcaagttccacatcgttcccaagttgagagttcagttatgtagagcagccgtgag 120  
 OY 2579 gaaactcgagactcgtggtggaagagaccttctgtaactttgaccacacacagcgtg 2638  
 DB 121 gaaactcgagactcgtggtggaagagaccttctgtaactttgaccacacacagcgtg 180  
 OY 2639 tgaagtcgaagaaatcccccactgtagatcgtatccacagcttaagcggcgccactc 2698  
 DB 181 tgaagtcgaagaaatcccccactgtagatcgtatccacagcttaagcggcgccactc 240  
 OY 2699 ccgattcgggttggaaagccgaaagactgaaatcccccacttaagcttgcctcatgcgc 2758  
 DB 241 ccgattcgggttggaaagccgaaagactgaaatcccccacttaagcttgcctcatgcgc 300  
 OY 2759 tggcttctaactctgataactcctcaagttatgtagtaagcgttggtagacagccgc 2818  
 DB 301 tggcttctaactctgataactcctcaagttatgtagtaagcgttggtagacagccgc 360

OY 2819 aactccataaaccttatactctcaccctggttacttactgactcgcgtacaggtatcat 2878  
 DB 361 aactccataaaccttatactctcaccctggttacttactgactcgcgtacaggtatcat 420  
 OY 2879 attaacagcaactcaaggagagctccctctgggagacttggcctgtaatatatgtctgc 2938  
 DB 421 attaacagcaactcaaggagagctccctctgggagacttggcctgtaatatatgtctgc 480  
 OY 2939 cttcgatcaagtaattccctggttctcaatgacagagccacaccccccgatgtctcgtgc 2998  
 DB 481 cttcgatcaagtaattccctggttctcaatgacagagccacaccccccgatgtctcgtgc 540  
 OY 2999 taagggtttaaagtttgcccaaggaccaccaataatgaaatattgtgaaatccctcag 3058  
 DB 541 taagggtttaaagtttgcccaaggaccaccaataatgaaatattgtgaaatccctcag 600  
 OY 3059 gatttctcttgcagaacatgtagcgtcgtacatacttcaatgtaggaattggaatgagcca 3118  
 DB 601 gatttctcttgcagaacatgtagcgtcgtacatacttcaatgtaggaattggaatgagcca 660  
 OY 3119 gttcttcagcaagacagagtaagttactcttctgtttaaacaatccctaccagttataatcaa 3178  
 DB 661 gttcttcagcaagacagagtaagttactcttctgtttaaacaatccctaccagttataatcaa 720  
 OY 3179 tttaattatgscatlggagatggaagatttggcaacagcgggtacaaaagatgtacga 3238  
 DB 721 tttaattatgscatlggagatggaagatttggcaacagcgggtacaaaagatgtacga 780  
 OY 3239 aataagcaataatagctgttactcgtttagacctagattcttaaaaatagttactc 3295  
 DB 781 aataagcaataatagctgttactcgtttagacctagattcttaaaaatagttactc 837

RESULT 10

ID AAC67032 standard; DNA; 6076 BP.

XX AAC67032;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 32.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000MO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI: 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -

PS Claim 16; Page 117-119; 14pp; English.

CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

XX Sequence 6076 BP, 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 22.5%; Score 746; DB 22; Length 6076;  
Best Local Similarity 98.7%; Pred. No. 1.6e-293;  
Matches 2196; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 31 tctgtcctagatattacaccactagaccacttttgccttgccttggaatgagagatcca 90  
Db 3577 tctgtcctagatattacaccactagaccacttttgccttgccttggaatgagagatcca 3636  
QY 91 ggtacggagaagaccgggagctcaccttggaccgagctgcccacaaagggttcaagaatccc 150  
Db 3637 ggtacggagaagaccgggagctcaccttggaccgagctgcccacaaagggttcaagaatccc 3696  
QY 151 ccgacatcttttgaggaagcccttacacaggagccctgcccacactttagatccaacacct 210  
Db 3697 ccgacatcttttgaggaagcccttacacaggagccctgcccacactttagatccaacacct 3756  
QY 211 caggtgacccctcctccagtaagtgatatacctgtctctgctggcgagagccacaaagagac 270  
Db 3757 caggtgacccctcctccagtaagtgatatacctgtctctgctggcgagagccacaaagagac 3816  
QY 271 tgccttagaagtgacgaagagcaactactgttggaattgtctgtaactaggtacagagccctc 330  
Db 3817 tgccttagaagtgacgaagagcaactactgttggaattgtctgtaactaggtacagagccctc 3876  
QY 331 gctaaagaagcccgagatttgcagagagaggaataacatacttgggttaccagtttcggggc 390  
Db 3877 gctaaagaagcccgagatttgcagagagaggaataacatacttgggttaccagtttcggggc 3936  
QY 391 gggcagcgatgtgctgacgagagcgacgagaaacactgagtcacagataccggcccacac 450  
Db 3937 gggcagcgatgtgctgacgagagcgacgagaaacactgagtcacagataccggcccacac 3996  
QY 451 acagccaaacaagtgtgagagagtttttgggagcagctggatttgcagagctgtgataccg 510  
Db 3997 acagccaaacaagtgtgagagagtttttgggagcagctggatttgcagagctgtgataccg 4056  
QY 511 ggggttgcgaccttagacgagcccccactacccggttaacccaagaagaaggggatcttc 570  
Db 4057 ggggttgcgaccttagacgagcccccactacccggttaacccaagaagaaggggatcttc 4116  
QY 571 tgggtcctcctgagcaccagaaggtcatttgatgtctatcaaaaaggccctgcgtgagcgacac 630  
Db 4117 tgggtcctcctgagcaccagaaggtcatttgatgtctatcaaaaaggccctgcgtgagcgacac 4176  
QY 631 gctctgacctcctgacgttaactaaaccttaaccttattgtgtgaatgaagcgtlaagga 690  
Db 4177 gctctgacctcctgacgttaactaaaccttaaccttattgtgtgaatgaagcgtlaagga 4236  
QY 691 gtagccggagagagttttaaccacaaaccttagacatagggagagacctgttgcctactcg 750  
Db 4237 gtagccggagagagttttaaccacaaaccttagacatagggagagacctgttgcctactcg 4296  
QY 751 tcaaaagaagctgtatccctgttagccaggtgttgcgccgtatgctcgtgaaggtatccgacgt 810  
Db 4297 tcaaaagaagctgtatccctgttagccaggtgttgcgccgtatgctcgtgaaggtatccgacgt 4356  
QY 811 gtggccatctgttcaagaagcgcttgacaatgtgaacttgggacagataataactgttaata 870  
Db 4357 gtggccatctgttcaagaagcgcttgacaatgtgaacttgggacagataataactgttaata 4416  
QY 871 gccccccatgcatgtgagaacatcgttcggcagccccccaaacgagatgagtaggacacggcc 930  
Db 4417 gccccccatgcatgtgagaacatcgttcggcagccccccaaacgagatgagtaggacacggcc 4476  
QY 931 cgaatgacccataccaagaactgtcttccacagaaaggttcaacttgcctccacagacc 990  
Db 4477 cgaatgacccataccaagaactgtcttccacagaaaggttcaacttgcctccacagacc 4536

QY 991 gcttcaaccctgccaactcttctgtcgtgaagagactgatgaaccagtgactcatgttgc 1050  
Db 4537 gcttcaaccctgccaactcttctgtcgtgaagagactgatgaaccagtgactcatgttgc 4596  
QY 1051 catcaactatgtattgaagagactgtgttcccgcaagagaccttaacagataccgtgact 1110  
Db 4597 catcaactatgtattgaagagactgtgttcccgcaagagaccttaacagataccgtgact 4656  
QY 1111 ggaagaagtgtcaacctgttctactgaacggaaagcagctatgtgtgtgaaggttaagagatg 1170  
Db 4657 ggaagaagtgtcaacctgttctactgaacggaaagcagctatgtgtgtgaaggttaagagatg 4716  
QY 1171 gctggggcgagctgtgtgagacgggaccccgacagctctggggccacacgcttgcgggaagga 1230  
Db 4717 gctggggcgagctgtgtgagacgggaccccgacagctctggggccacacgcttgcgggaagga 4776  
QY 1231 acttcaagcgcaaaaggctgagcttcaatgacctcaccgcaaaacttgcggctgcccgaagg 1290  
Db 4777 acttcaagcgcaaaaggctgagcttcaatgacctcaccgcaaaacttgcggctgcccgaagg 4836  
QY 1291 aaatccataacattatacaggaacagagatagcttgcgtactgcacagtlacacgg 1350  
Db 4837 aaatccataacattatacaggaacagagatagcttgcgtactgcacagtlacacgg 4896  
QY 1351 gcaatctataacaaagggttgccttactcctcagcgagggggaaataaagaacaaagg 1410  
Db 4897 gcaatctataacaaagggttgccttactcctcagcgagggggaaataaagaacaaagg 4956  
QY 1411 gaaattctaaagcctatagaagccttacaatttgcacaaaggcttagctataatacctgt 1470  
Db 4957 gaaattctaaagcctatagaagccttacaatttgcacaaaggcttagctataatacctgt 5016  
QY 1471 ccttgagatacagaagaagccaaagatccatatactagaagggaacagatgctgcagggtt 1530  
Db 5017 ccttgagatacagaagaagccaaagatccatatactagaagggaacagatgctgcagggtt 5076  
QY 1531 gccaaagcagcagcccgagctgtttaaactcttccctataatagaacgcccacaaagccca 1590  
Db 5077 gccaaagcagcagcccgagctgtttaaactcttccctataatagaacgcccacaaagccca 5136  
QY 1591 gaaaccaagcagcagtaacaccttagaagactggcagaagataaanaagatagaccagttc 1650  
Db 5137 gaaaccaagcagcagtaacaccttagaagactggcagaagataaanaagatagaccagttc 5196  
QY 1651 tctgaagactccggaaggagacctgtctactactataatgggaaggaatacctgcgccacaa 1710  
Db 5197 tctgaagactccggaaggagacctgtctactactataatgggaaggaatacctgcgccacaa 5256  
QY 1711 gaaagggtagaatatgtccaacagatatacgtcttaaccacacttagaagactaaacacgtg 1770  
Db 5257 gaaagggtagaatatgtccaacagatatacgtcttaaccacacttagaagactaaacacgtg 5316  
QY 1771 cagcagttgttcaagaacatcccttatacattgttcttgaggttaccaggaatgtgctgactcg 1830  
Db 5317 cagcagttgttcaagaacatcccttatacattgttcttgaggttaccaggaatgtgctgactcg 5376  
QY 1831 gtgtgtcaaacattgtgtgtgccttgcggccaagctgtgttaatgtcctaactccgaatatacc 1890  
Db 5377 gtgtgtcaaacattgtgtgtgccttgcggccaagctgtgttaatgtcctaactccgaatatacc 5436  
QY 1891 ggaagaagactaaggggaagccacacagcgctcactgtggaagtgtgacttcaactgaagta 1950  
Db 5437 ggaagaagactaaggggaagccacacagcgctcactgtggaagtgtgacttcaactgaagta 5496  
QY 1951 aagccggctaaatacggaaacaataatctatgtgttctttagacacacttccaagatg 2010  
Db 5497 aagccggctaaatacggaaacaataatctatgtgttctttagacacacttccaagatg 5556  
QY 2011 gtagaggtataccttaacagaagaagacttcaacacgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2070  
Db 5557 gtagaggtataccttaacagaagaagacttcaacacgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5616  
QY 2071 gaaattttccaagaattgtgaatatcctaaggtlaataggttcagacaaatgttccacttcc 2130

Db	5617	gaatttttccaagatttggaatacccaaggtaataaggtccagcaatggtccagcttc	5676
QY	2131	gtgtcccaagtaaatcaggagctggccaagatatgtggagaltgtggaactgcatgt	2190
Db	5677	gtgtcccaagtaagtcaggagctggccaagatatgtggagaltgtggaactgcatgt	5736
QY	2191	gcatacagagcccaagctcagacaggtagaagagatgaataacatlaagaagacc	2250
Db	5737	gcatacagagcccaagctcagacaggtagaagagatgaataacatlaagaagacc	5796
QY	2251	cttac 2255	
Db	5797	cttac 5801	
RESULT 11			
XX	AAAT74884	standard; cDNA: 7892 BP.	
XX	AAAT74884;		
XX	09-FEB-1998	(first entry)	
XX	De	Miniature swine retrovirus cDNA.	
XX	Retrovirus; porcine; GAG protein; POL protein; ENV protein;		
KW	xenotransplantation; infectious; provirus; organ transplant; donor;		
KX	activated virus; PCR; ss.		
OS	Porcine retrovirus.		
XX	Key	Location/Qualifiers	
EH	CDS	585..2159	
FT		/*tag= a	
FT	mat_peptide	585..2156	
FT		/*tag= b	
FT	CDS	/*note= "putative GAG protein"	
FT		2307..5744	
FT		/*tag= c	
FT	mat_peptide	2307..5741	
FT		/*tag= d	
FT	CD9	/*note= "putative POL protein"	
FT		5620..7536	
FT		/*tag= e	
FT	mat_peptide	5620..7533	
FT		/*tag= f	
FT		/*note= "putative ENV protein"	
XX	WO9721836-A1.		
XX	19-JUN-1997.		
XX	PF	13-DEC-1996; 96WO-US19680.	
XX	PR	14-DEC-1995; 95US-0572645.	
XX	PA	(GEHO ) GEN HOSPITAL CORP.	
XX	PI	Fishman JA;	
XX	WPI; 1997-332804/30.		
DR	P-PSDB; AAW32096-W32098.		
PT	New nucleic acid from porcine retroviruses - used for detecting		
PT	viruses in transplant or other tissue and for assessing risk of		
PT	transmitting infection to graft recipient		
XX	Claim 22; Fig 3; 128bp; English.		
CC	This cDNA sequence represents a porcine retrovirus from miniature swine		
CC	containing the putative coding regions for viral GAG, POL, and ENV		
CC	proteins. This sequence and PCR fragments generated from it		

CC (see AATV4812-774882) can be used to screen organs for the presence of  
CC porcine retroviruses prior to xenotransplantation. Transplantation can  
CC increase the likelihood of retroviral activation if intact and  
CC infectious proviruses are present. The porcine retroviral sequence can be  
CC used to generate probes to determine the level (e.g. copy number) of  
CC strain (i.e. potentially replicating) porcine provirus sequences in a  
CC extract of xenograft transplantation donors. It can be used to detect  
CC mutations, genetic lesions or viral recombinants and also to determine  
CC the histological localisation of activated retroviruses. Using Polymerase  
CC Chain Reaction DNA Quantitation (PQD) on blood mononuclear cells,  
CC infectivity titration and susceptibility testing can be performed.  
CC ultimately animal donors without intact porcine retroviral sequences or a  
CC lower copy number of viral elements could be selected.  
XX  
SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match	21.2%	Score	704:	DB	18:	Length	7892:
Best Local Similarity	98.7%:	Pred. No.	1.7e-276:				
Matches	2204:	Conservative	0:	Mismatches	30:	Indels	0:

QY	22	gatgctctctctgtgcttgatatacaaccctatgccaacacatttgctctgaaatg	81
Db	2970	gatgctctctctgtgcttgatatacaaccctatgccaacacatttgctctgaaatg	3029
QY	82	aaagatccagtgtaacgggaagacccgggacgctcaactgtgaccgagtgcccaaggttc	141
Db	3030	aaagatccagtgtaacgggaagacccgggacgctcaactgtgaccgagtgcccaaggttc	3089
QY	142	aagaatctcccgacaacctctcttgacgaagccctacaacagagagacttggccaaattcaagatc	201
Db	3090	aagaatctcccgacaacctctcttgacgaagccctacaacagagagacttggccaaattcaagatc	3149
QY	202	caacacctcaagtgtaacctctctccagtcacgtcgtgatgactgctctcttgccgggaagccac	261
Db	3150	caacacctcaagtgtaacctctctccagtcacgtcgtgatgactgctctcttgccgggaagccac	3209
QY	262	aacacagatgctgtataaagttacgaagacacactcgtctgaaattgtctacataaggtac	321
Db	3210	aacacagatgctgtataaagttacgaagacacactcgtctgaaattgtctacataaggtac	3269
QY	322	agaagctctgtctaagaagaagcccaagatttgcaaggagaggtaaacatacttgggtacagt	381
Db	3270	agaagctctgtctaagaagaagcccaagatttgcaaggagaggtaaacatacttgggtacagt	3329
QY	382	tgcgcggcgcgagcagcagatgctgcgaaggggacacggaaagaaaccttattccagatcacg	441
Db	3330	tgcgcggcgcgagcagcagatgctgcgaaggggacacggaaagaaaccttattccagatcacg	3389
QY	442	gccccaacacacacgaacaaacagtgacagagttcttctgggacacagctgatttgcagaagt	501
Db	3390	gccccaacacacacgaacaaacagtgacagagttcttctgggacacagctgatttgcagaagt	3449
QY	502	tggatcccggggtttgcgaacttagcagcccaactctaccgcttaaccgaagaaaaagg	561
Db	3450	tggatcccggggtttgcgaacttagcagcccaactctaccgcttaaccgaagaaaaagg	3509
QY	562	ggattctccgggctctcttagcacaacgaagacatttgatgcta tcaaaaagggccgtctg	621
Db	3510	ggattctccgggctctcttagcacaacgaagacatttgatgcta tcaaaaagggccgtctg	3569
QY	622	agcgacactgtcttgccctccctgacgttaactaaaccttacccttattgtgatag	681
Db	3570	agcgacactgtcttgccctccctgacgttaactaaaccttacccttattgtgatag	3629
QY	682	cgttaagggagtagcccgaggagttttaacccaaccttaggaccatggagagaacctgtc	741
Db	3630	cgttaagggagtagcccgaggagttttaacccaaccttaggaccatggagagaacctgtc	3689
QY	742	gcttactctgtcaaaagagcttgatccctgttagccagtggttggccggtatcttgaagct	801
Db	3690	gcttactctgtcaaaagagcttgatccctgttagccagtggttggcccatatgctcctaagct	3749

QY 802 atcgcaactgtgcccatactgctcaagagcgtgcacaattgacttgggacagaatata 861  
 Db 3750 atcgcaactgtgcccatactgctcaagagcgtgcacaattgacttgggacagaatata 3809  
 QY 862 acttaataagcccccatgcatctggaagaacatgcttcggaagccccacagatgag 921  
 Db 3810 actgtaataagcccccatgcatctggaagaacatgcttcggaagccccacagatgag 3869  
 QY 922 accaagcccgcaatgaccccaatacaaaagctgtctctcaagaaggggttcaatttcgct 981  
 Db 3870 accaagcccgcaatgaccccaatacaaaagctgtctctcaagaaggggttcaatttcgct 3929  
 QY 982 ccaacagccgctctcaacctgtccactcttctgctgaagagatgataaacaagtact 1041  
 Db 3930 ccaacagccgctctcaacctgtccactcttctgctgaagagatgataaacaagtact 3989  
 QY 1042 catgattgcatcaactatattgatttgaggagactggggttcggaagagacttacaagata 1101  
 Db 3990 catgattgcatcaactatattgatttgaggagactggggttcggaagagacttacaagata 4049  
 QY 1102 ccgctgactggaagaatgctaaacctgttcaactgacggaagcagctatgtgtggaagt 1161  
 Db 4050 ccgctgactggaagaatgctaaacctgttcaactgacggaagcagctatgtgtggaagt 4109  
 QY 1162 aagagagatgctgctgggagcagctgtgtggaagcagccgacagatctgggccaagcctg 1221  
 Db 4110 aagagagatgctgctgggagcagctgtgtggaagcagccgacagatctgggccaagcctg 4169  
 QY 1222 ccggaagagacttcaacgcaaaagctgagctcagtgccctcaagaagcttgcgctg 1281  
 Db 4170 ccggaagagacttcaacgcaaaagctgagctcagtgccctcaagaagcttgcgctg 4229  
 QY 1282 gccgaaggggaatcctaataacattatacgaagcaggtatgcttctgcaactgacac 1341  
 Db 4230 gccgaaggggaatcctaataacattatacgaagcaggtatgcttctgcaactgacac 4289  
 QY 1342 gfaccgaggccatctataaacaagggtgtgttacttcaagcagggagggaaataag 1401  
 Db 4290 gfaccgaggccatctataaacaagggtgtgttacttcaagcagggagggaaataag 4349  
 QY 1402 aacaagaagggaatcctaagctctataagaagctctacatttgcgaagaagcttactatt 1461  
 Db 4350 aacaagaagggaatcctaagctctataagaagctctacatttgcgaagaagcttactatt 4409  
 QY 1462 ataactgtctggaacatcaagaagcacaagatctcatatcagaaggaaccagatgct 1521  
 Db 4410 ataactgtctggaacatcaagaagcacaagatctcatatcagaaggaaccagatgct 4469  
 QY 1522 gaccggtgtgccaagcagcagccagctgttaaccttgcctctataatagaacgccc 1581  
 Db 4470 gaccggtgtgccaagcagcagccagctgttaaccttgcctctataatagaacgccc 4529  
 QY 1582 aaagcccaagaaccagacagctacacccctagaagaagctgcgaagatataaaaaata 1641  
 Db 4530 aaagcccaagaaccagacagctacacccctagaagaagctgcgaagatataaaaaata 4589  
 QY 1642 gaccagttctctgagactccggaaggagactgctatacctcaatatygaagaaatcctg 1701  
 Db 4590 gaccagttctctgagactccggaaggagactgctatacctcaatatygaagaaatcctg 4649  
 QY 1702 ccccaacaagaagggttagaatatgtccaacagatactgttcaacccactgagact 1761  
 Db 4650 ccccaacaagaagggttagaatatgtccaacagatactgttcaacccactgagact 4709  
 QY 1762 aaacactgcagcagctgtgtcagaacatccctctatcatgttctgaggtaccagagatg 1821  
 Db 4710 aaacactgcagcagctgtgtcagaacatccctctatcatgttctgaggtaccagagatg 4769  
 QY 1822 gctgactcgtgtgccaacatgtgttcctcgcagctgtgttaattgttaactcctccaga 1881  
 Db 4770 gctgactcgtgtgccaacatgtgttcctcgcagctgtgttaattgttaactcctccaga 4829  
 QY 1882 ataccctcaggaagagactlaagggaagcaccagcgctcaactggaagtgagacttc 1941

Db 4830 atgcccacaggggaagagactlaagggaagccacccagcgctcactggaagtgagactc 4889  
 QY 1942 actgaggtlaaagccggtctaataatcaggaacaataatactatgtgtttgtgacacctt 2001  
 Db 4890 actgaggtlaaagccggtctaataatcaggaacaataatactatgtgtttgtgacacctt 4949  
 QY 2002 tcaagatgtgtagaggttatctctactaagaagaagacttcaacgctgtgtgtaagaa 2061  
 Db 4950 tcaagatgtgtagaggttatctctactaagaagaagacttcaacgctgtgtgtaagaa 5009  
 QY 2062 atactggaaggaatttcccaagatttgaatacctaaggttaaggttcaagaactgt 2121  
 Db 5010 atactggaaggaatttcccaagatttgaatacctaaggttaaggttcaagaactgt 5069  
 QY 2122 ccagcttctgtgcccaagtagtaagtcagagctgcccagaatattgaggatgttgaaa 2181  
 Db 5070 ccagcttctgtgcccaagtagtaagtcagagctgcccagaatattgaggatgttgaaa 5129  
 QY 2182 ctgcatgtgatatagaagccccaagctcaggaacaggttaagagagatagaataacatt 2241  
 Db 5130 ctgcatgtgatatagaagccccaagctcaggaacaggttaagagagatagaataacatt 5189  
 QY 2242 aaagagacccttac 2255  
 Db 5190 aaagagacccttac 5203

## RESULT 12

AAF77727

ID AAF77727 standard; DNA: 8132 BP.

XX AAF77727;

DT 23-MAY-2001 (first entry)

DE Nucleotide sequence of a retrovirus found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine; KW ss.

XX Unidentified.

PN US6190861-B1.

PD 20-FEB-2001.

XX 13-DEC-1996; 9605-0766528.

XX 14-DEC-1995; 95US-0572645.

PA (GENO ) GEN HOSPITAL CORP.

XX Fishman JA;

DR WPI: 2001-256211/26.

DR P-PsDB: AAB73285, AAB73286, AAB73287.

PT Assessing risk of endogenous retroviruses in clinical practice and in

XX xenotransplantation, comprises using probe sequences derived from swine

XX or miniature swine retroviral genome -

XX Claim 1; Fig 3; 127pp; English.

The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid (a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible

in the outbred human organ donor population and is important to the development of human xenotransplantation.

CC Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

Query Match: 21.2%; Score 704; DB 22; Length 8132;

Best Local Similarity 98.7%; Pred. No. 1.7e-276; Mismatches 30; Indels 0; Gaps 0;

Matches 2204; Conservative

22 gatgacctctctgctgagatatacaccacatagcaaacattttgcttcgaatg 81  
2970 gatgacctctctgctgagatatacaccacatagcaaacattttgcttcgaatg 3029  
82 agagatccaggtacggaagaacccgggacagctacttgaccctgcccgaaggttc 141  
3030 agagatccaggtacggaagaacccgggacagctacttgaccctgcccgaaggttc 3089  
142 aagaactcccgacacatttgacgaagccctacacagagacgtggccaaattcaagatc 201  
3090 aagaactcccgacacatttgacgaagccctacacagagacgtggccaaattcaagatc 3149  
202 caaacacctcaggttgaacctctccagtagcgttgatgaacctctctgycgggagccacc 261  
3150 caaacacctcaggttgaacctctccagtagcgttgatgaacctctctgycgggagccacc 3209  
262 aaaaaggaactgttagaaggtacgaagacacatctgctggaattgtgtgacgttagc 321  
3210 aaaaaggaactgttagaaggtacgaagacacatctgctggaattgtgtgacgttagc 3269  
322 aagaacctctgctgaagaagccacagatttgacgaagagaggttaacatacttggggtagc 381  
3270 aagaacctctgctgaagaagccacagatttgacgaagagaggttaacatacttggggtagc 3329  
382 ttgcygggcygggacgagatggtctgaacggagccacggaagaacactgtatgccagatccg 441  
3330 ttgcygggcygggacgagatggtctgaacggagccacggaagaacactgtatgccagatccg 3389  
442 gcccaaacacagacgaacaaagtagagagatttttggggacagcttgatcttcagagatg 501  
3390 gcccaaacacagacgaacaaagtagagagatttttggggacagcttgatcttcagagatg 3449  
502 tgaatcccggggttttgcgaccttgaacgacccacctacccgcttaaccnaagaagaagg 561  
3450 tgaatcccggggttttgcgaccttgaacgacccacctacccgcttaaccnaagaagaagg 3509  
562 ggatcttctctggtctctctgagcaccagaagacatttgatctatacaaaaggccctctg 621  
3510 ggatcttctctggtctctctgagcaccagaagacatttgatctatacaaaaggccctctg 3569  
622 agcgacacctgctcgtgacctctccgtgacgttaacacacatttaccctttagtggatgag 681  
3570 agcgacacctgctcgtgacctctccgtgacgttaacacacatttaccctttagtggatgag 3629  
682 cgtlaaggaatagcccgagagattttaaaccnaacccctaggaaccaatggaagaagccgttc 741  
3630 cgtlaaggaatagcccgagagattttaaaccnaacccctaggaaccaatggaagaagccgttc 3669  
742 gcttaacctgttaagaagacgtgacacctgttagcaggtgttgccgcatgtctaaagct 801  
3690 gcttaacctgtgttaagaagacgtgacacctgttagcaggtgttgccgcatgtctaaagct 3749  
802 atcgagactgtgacatctgttcaagaagcgttacaacatgtgactttgggacagaatata 861  
3750 atcgagactgtgacatctgttcaagaagcgttacaacatgtgactttgggacagaatata 3809  
862 actgtaataagcccccacatgtcatgttggagaaacttggtggcggcccccagaccgtagtg 921  
3810 actgtaataagcccccacatgtcatgttggagaaacttggtggcggcccccagaccgtagtg 3869  
922 accaagcccgcatgacacacatcaaaagcgtgtcttccacagagaggttcaacttcgct 981  
3870 accaagcccgcatgacacacatcaaaagcgtgtcttccacagagaggttcaacttcgct 3929

982 ccacacgacctctcaaacctcgacactctctgctgaagaagactgataaacagtagct 1041  
3930 ccacacgacctctcaaacctcgacactctctgctgaagaagactgataaacagtagct 3989  
1042 catgattgcatcaactatgattgaagaagacgtgggtccgcaagaaccttacaagata 1101  
3990 catgattgcatcaactatgattgaagaagacgtgggtccgcaagaaccttacaagata 4049  
1102 ccgctgacttggaaggtgtcaacctgttcaactacgaagaacacatgtgtgtgaaggt 1161  
4050 ccgctgacttggaaggtgtcaacctgttcaactacgaagaacacatgtgtgtgaaggt 4109  
1162 aagaagatggtctggtggtcgagtagtggtagcggagacccgacagatctgggccaagcctg 1221  
4110 aagaagatggtctggtggtcgagtagtggtagcggagacccgacagatctgggccaagcctg 4169  
1222 ccggaagaagacttcaagcccaaaagctgagctcagtcagccccaagacgtcttcggtctg 1281  
4170 ccggaagaagacttcaagcccaaaagctgagctcagtcagccccaagacgtcttcggtctg 4229  
1282 gccgaagggaattccataaacaattatatacgaacagtagtaccttgcgactgcacac 1341  
4230 gccgaagggaattccataaacaattatatacgaacagtagtaccttgcgactgcacac 4289  
1342 gtacacggtggccatctataaacaaggggttgccttaacctcagcagggagggaaataag 1401  
4290 gtacatggggccatctataaacaaggggttgccttaacctcagcagggagggaaataag 4349  
1402 aacaaagaagaattcttaagccttatagaagccttaccatttgcacaaaagcctagctat 1461  
4350 aacaaagaagaattcttaagccttatagaagccttaccatttgcacaaaagcctagctat 4409  
1462 ataacatgctcctggaacatagaagaacccaagaatcccatatctagaagggaaccagatg 1521  
4410 ataacatgctcctggaacatagaagaacccaagaatcccatatctagaagggaaccagatg 4469  
1522 gaccgggttgcgaacagcagccagcgtgttgaaccttgcctataatgaagaacgccc 1581  
4470 gaccgggttgcgaacagcagccagcgtgttgaaccttgcctataatgaagaacgccc 4529  
1582 aagaagcccaagacccgaagcagacagtagtaacacccctagagagacttgcgaagaagata 1641  
4530 aagaagcccaagacccgaagcagacagtagtaacacccctagagagacttgcgaagaagata 4589  
1642 gaccagttctctgagacctccggaaggacgtgtataacctcaatggtgaagaagaactctg 1701  
4590 gaccagttctctgagacctccggaaggacgtgtataacctcaatggtgaagaagaactctg 4649  
1702 ccccaaaagaaggttgaataatgtccaaagatacatcgtctaaacccactagaagact 1761  
4650 ccccaaaagaaggttgaataatgtccaaagatacatcgtctaaacccactagaagact 4709  
1762 aaacacctgacaggttgcgtcaagaacatccctatcatatgtcttgagctacagagatg 1821  
4710 aaacacctgacaggttgcgtcaagaacatccctatcatatgtcttgagctacagagatg 4769  
1822 gctgactcgtgtgtcaacaattgtgtccctgcagcgtgttaatgtataatcccttcaga 1881  
4770 gctgactcgtgtgtcaacaattgtgtccctgcagcgtgttaatgtataatcccttcaga 4829  
1882 ataccctccgggaagaagactaaagggaagccacccagcggtcaacgtggaagtgtgacttc 1941  
4830 ataccctccgggaagaagactaaagggaagccacccagcggtcaacgtggaagtgtgacttc 4889  
1942 actgaggttaagcccgctaaatlaacggaacaaatatactatltgttttggtagaacctt 2001  
4890 actgaggttaagcccgctaaatlaacggaacaaatatactatltgttttggtagaacctt 4949  
2002 tcaagatggttaaggtcttactactaagaagaagacttcaacggtgtgtgtagaagaa 2061  
4950 tcaagatggtgttagaggttactactactaagaagaagacttcaacggtgtgtgtagaagaa 5009



QY 2062 atactgagagaatttttccaagatttggaaactaactaagtataggttcagacaatggt 2121  
|||||  
Db 5010 atactgagagaatttttccaagatttggaaactaactaagtataggttcagacaatggt 5069  
QY 2122 ccagcttctggtccagagtaagtcagagactgagccagatactggagatctatggaaa 2181  
|||||  
Db 5070 ccagcttctggtccagagtaagtcagagactgagccagatactggagatctatggaaa 5129  
QY 2182 ctgactgtgcatatagacaccccaagctcagagacaggttagagagatgaataaccatc 2241  
|||||  
Db 5130 ctgactgtgcatatagacaccccaagctcagagacaggttagagagatgaataaccatc 5189  
QY 2242 aaagagacccttac 2255  
|||||  
Db 5190 aaagagacccttac 5203

RESULT 13  
AAC67022  
ID AAC67022 standard; DNA; 4918 BP.  
XX  
AC AAC67022;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE PERV env protein coding sequence SEQ ID NO: 22.  
XX  
KW Xenotransplantation; infectious agent; vaccine; ds.  
XX  
OS Porcine endogenous retrovirus.  
XX  
PN WO200071726-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 24-MAY-2000; 2000WO-US14296.  
XX  
PR 24-MAY-1999; 99US-0135631.  
XX  
PA (MAYO-) MAYO MEDICAL VENTURES.  
XX  
PI Federspiel MJ;  
XX  
DR MPI; 2001-032041/04.  
XX  
PT Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
XX  
PS Claim 16; Page 109-111; 144pp; English.  
XX  
CC The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX  
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 17.2%; Score 570; DB 22; Length 4918;  
Best Local Similarity 98.7%; Pred. No. 4.4e-222;  
Matches 1670; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 684 taaggagtagcccgaggaggttttaaccacaacccctagggaccatggagagactgtgac 743  
|||||  
Db 199 taaggagtagcccgaggaggttttaaccacaacccctagggaccatggagagactgtgac 258  
QY 744 ctactgtcacaagaagcttgaacctgttagccagtggttggccgtatgtctgaagagctat 803  
|||||  
Db 259 ctactgtcacaagaagcttgaacctgttagccagtggttggccgtatgtctgaagagctat 318  
QY 804 cgcagctgtgagcactgtgtcagaagagcgtcgaacaattgacttggagacgaataaac 863  
|||||  
Db 319 cgcagctgtgagcactgtgtcagaagagcgtcgaacaattgacttggagacgaataaac 378  
QY 864 tgtatagcccccatgcatgttggaagaatggttcgcgagcccccagaccgatgtgac 923  
|||||  
Db 379 tgtatagcccccatgcatgttggaagaatggttcgcgagcccccagaccgatgtgac 438  
QY 924 caacgcccgcatagcacacatacacaagctgtcttcacagagaggttcaacttcgctcc 983  
|||||  
Db 439 caacgcccgcatagcacacatacacaagctgtcttcacagagaggttcaacttcgctcc 498  
QY 984 accagccgctctcaacctgcccactcttctgcctggaagagctatgaaccagtgactca 1043  
|||||  
Db 499 accagccgctctcaacctgcccactcttctgcctggaagagctatgaaccagtgactca 558  
QY 1044 tgattgcatcaactatgtatgtgagagactgggttcgcgacaagaccttacaagacatacc 1103  
|||||  
Db 559 tgattgcatcaactatgtatgtgagagactgggttcgcgacaagaccttacaagacatacc 618  
QY 1104 gctgactgagagaagtgtaacctgttcaactgtacagcgaagcactatgtgtggaagttaa 1163  
|||||  
Db 619 gctgactgagagaagtgtaacctgttcaactgtacagcgaagcactatgtgtggaagttaa 678  
QY 1164 gaggaatgctggggcgaggtgtggaagagccgcagatctggtgcccagacgctgccc 1223  
|||||  
Db 679 gaggaatgctggggcgaggtgtggaagagccgcagatctggtgcccagacgctgccc 738  
QY 1224 ggaaggaacttcagcgcacaagagctgagctatgctccacgcgaagcttggcgctggc 1283  
|||||  
Db 739 ggaaggaacttcagcgcacaagagctgagctatgctccacgcgaagcttggcgctggc 798  
QY 1284 cgaagggaaatccataaacaattatagcagacagatgacttgcgactgcacagct 1343  
|||||  
Db 799 cgaagggaaatccataaacaattatagcagacagatgacttgcgactgcacagct 858  
QY 1344 acacggggccatctataacaagaggtgtgttaacctcagcagcaggaaggaataaagaa 1403  
|||||  
Db 859 acacggggccatctataacaagaggtgtgttaacctcagcagcaggaaggaataaagaa 918  
QY 1404 caaagagaatcttaagcctataggaagccttaacttggccaaaaggtagctattat 1463  
|||||  
Db 919 caaagagaatcttaagcctataggaagccttaacttggccaaaaggtagctattat 978  
QY 1464 acaactgtctcgacatcagaagaagccaaagatcctatctagaggaaccgagtgcga 1523  
|||||  
Db 979 acaactgtctcgacatcagaagaagccaaagatcctatctagaggaaccgagtgcga 1038  
QY 1524 ccgggttccaaagcagcagcccgagctgttaacctctgcctataataggaagcgcacaa 1583  
|||||  
Db 1039 ccgggttccaaagcagcagcccgagctgttaacctctgcctataataggaagcgcacaa 1098  
QY 1584 agccccaagaaccagacagacaglacacctagaagactgtgcaagagataaaagataga 1643  
|||||  
Db 1099 agccccaagaaccagacagacaglacacctagaagactgtgcaagagataaaagataga 1158  
QY 1644 ccagttctctgagactccggagggagactgtctatataccatctatggaggaatctctg 1703  
|||||  
Db 1159 ccagttctctgagactccggagggagactgtctatataccatctatggaggaatctctg 1218  
QY 1704 ccacaagaagaggttagaataatgtccacaagatacatgtcttaaacccactaggaactaa 1763  
|||||  
Db 1219 ccacaagaagaggttagaataatgtccacaagatacatgtcttaaacccactaggaactaa 1278  
QY 1764 acaactgacagcagttgtgtcagaacatcccttatcatgttctgaggtaccagagatggc 1823



DB 541 tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgcatcgttgaac 600

RESULT 15  
AAH26258  
ID AAH26258 standard; DNA; 599 BP.  
XX  
XX AAH26258;  
AC  
XX 02-OCT-2001 (first entry)  
DT  
XX Pig endogenous retrovirus clone 11619-1.  
DE  
XX PERV; pig; xenotransplantation; ss.  
KM  
XX PERV; pig endogenous retrovirus.  
OS  
XX Wo200153825-A1.  
PN  
XX 26-JUL-2001.  
PD  
XX 19-JAN-2001; 2001WO-US01857.  
PE  
XX 19-JAN-2000; 2000US-0177003.  
PR 16-FEB-2000; 2000US-0182965.  
PR 27-OCT-2000; 2000US-0243695.  
XX  
XX (BIOT-) BIOTRANSPLANT INC.  
PA  
XX Patience C;  
PI  
XX WPI; 2001-483157/52.  
DR  
XX  
XX New inbred swine of DD haplotype defective for transmission of porcine  
PT endogenous retroviruses, useful as source of clinical organs, tissues  
PT and cells, or as source of stem cells for xenogenic stem cell and  
PT thymic replacement therapy  
PT  
XX  
XX Example 4; Fig 5; 66pp; English.  
PS  
XX The present sequence is that of pig endogenous retrovirus (PERV)  
CC PCR-amplified clone 11619-1. Peripheral blood mononuclear cells  
CC from swine leukocyte antigen inbred miniature swine number 11619  
CC or 12002 were induced for PERV expression, and then co-cultured  
CC with uninfected human 293 cells. The infection of the cells was  
CC determined by the presence of reverse transcriptase (RT) activity  
CC in the culture supernatants. Genomic DNA was isolated and PCR  
CC products were generated using LTR-to-LTR PCR from cultures showing  
CC high levels of RT activity. The PCR products were cloned and  
CC sequenced to determine the different types of replication competent  
CC PERV present. 9 PERV sequences were derived from pig 11619 (see  
CC AAH26258-66) and 7 from pig 12002 (see AAH26251-57). Having  
CC determined the spectrum of sequences capable of replicating in  
CC human cells it is then possible to determine if these sequences are  
CC present in the genome of D/D haplotype miniature swine. Their  
CC absence then shows a lack of production of human-tropic virus from  
CC the cells of the D/D haplotype animals. The object of the invention  
CC is to provide a swine breed which does not produce PERV capable of  
CC infecting human cells, as measured by high-stringency in vitro  
CC assays. Such animals, which are preferably miniature pigs of the  
CC DD haplotype, have a unique advantage over other swine breeds with  
CC respect to microbiological safety for xenotransplantation. They  
CC are a valuable source of organs, tissues and cells, including a  
CC source of stem cells for xenogenic stem cell and thymic replacement  
CC therapy, e.g. in treating AIDS cases. Methods of producing such  
CC animals and for screening animals for the desired properties are  
CC disclosed.  
XX  
XX Sequence 599 BP; 166 A; 148 C; 129 G; 156 T; 0 other;

Query Match 15.0%; Score 497; DB 22; Length 599;  
Best Local Similarity 99.7%; Pred. No. 2.6e-192;

	Matches	597;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	2669	atgcatcccaagtttaagcgcgcacacccctccgattcgggtggaagccgaaagactg								2728
DB	1	atgcatcccaagtttaagcgcgcacacccctccgattcgggtggaagccgaaagactg								60
QY	2729	aaatcccttaagcttcgcctccatccgctggttcttaactctgtcaataactctcaa								2788
DB	61	aaatcccttaagcttcgcctccatccgctggttcttaactctgtcaataactctcaa								120
QY	2789	gttaatggttaagccttctggaacagccgaactccataaacccttaactcactctg								2848
DB	121	gttaatggttaagccttctggaacagccgaactccataaacccttaactcactctg								180
QY	2849	ttactactgactccggttaaggtatataataaagaactaagggaggtcccttg								2908
DB	181	ttactactgactccggttaaggtatataataaagaactaagggaggtcccttg								240
QY	2909	gggacctgtggtcctgaatataatgtctgccttcgaactgaactccctgtctcaatgac								2968
DB	241	gggacctgtggtcctgaatataatgtctgccttcgaactgaactccctgtctcaatgac								300
QY	2969	caagccacaccccccgatgactcgttgcttaacggttttaacgtttgcccagaccacca								3028
DB	301	caagccacaccccccgatgactcgttgcttaacggttttaacgtttgcccagaccacca								360
QY	3029	aataatgaagaataatgtggaatcctcagagattccttgcgaagaatgagctgcata								3088
DB	361	aataatgaagaataatgtggaatcctcagagattccttgcgaagaatgagctgcata								420
QY	3089	actcctaagatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaat								3148
DB	421	actcctaagatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaat								480
QY	3149	ttgtttaacaatcctacccaatttaatttaatttaatttaatttaatttaatttaatttaatt								3208
DB	481	ttgtttaacaatcctacccaatttaatttaatttaatttaatttaatttaatttaatttaatt								540
QY	3209	tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgtcatcgttgaac								3267
DB	541	tggcaacgcyggtacaaaagatgacgaataaagcaataaagctgtcatcgttgaac								599

Search completed: February 24, 2002, 07:34:23  
Job time: 16368 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      February 24, 2002, 07:42:18 ; Search time 584.4 Seconds
              (without alignments)
              4336.505 Million cell updates/sec
```

```

Title: US-09-171-553B-9
Perfect score: 2956
Sequence: 1 tgccttttaggttagaac.....aaaaaaaaaaaaaaaaaa 2956

```

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

word size :

Total number of hits satisfying chosen parameters: 1861242

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database :

1: N\_genseq\_1101.\*  
2: /SID52/gcgdata/genseq/genseqn/NA1980.DAT.\*  
3: /SID52/gcgdata/genseq/genseqn/NA1981.DAT.\*  
4: /SID52/gcgdata/genseq/genseqn/NA1982.DAT.\*  
5: /SID52/gcgdata/genseq/genseqn/NA1983.DAT.\*  
6: /SID52/gcgdata/genseq/genseqn/NA1984.DAT.\*  
7: /SID52/gcgdata/genseq/genseqn/NA1985.DAT.\*  
8: /SID52/gcgdata/genseq/genseqn/NA1986.DAT.\*  
9: /SID52/gcgdata/genseq/genseqn/NA1987.DAT.\*  
10: /SID52/gcgdata/genseq/genseqn/NA1988.DAT.\*  
11: /SID52/gcgdata/genseq/genseqn/NA1989.DAT.\*  
12: /SID52/gcgdata/genseq/genseqn/NA1990.DAT.\*  
13: /SID52/gcgdata/genseq/genseqn/NA1991.DAT.\*  
14: /SID52/gcgdata/genseq/genseqn/NA1992.DAT.\*  
15: /SID52/gcgdata/genseq/genseqn/NA1993.DAT.\*  
16: /SID52/gcgdata/genseq/genseqn/NA1994.DAT.\*  
17: /SID52/gcgdata/genseq/genseqn/NA1995.DAT.\*  
18: /SID52/gcgdata/genseq/genseqn/NA1996.DAT.\*  
19: /SID52/gcgdata/genseq/genseqn/NA1997.DAT.\*  
20: /SID52/gcgdata/genseq/genseqn/NA1998.DAT.\*  
21: /SID52/gcgdata/genseq/genseqn/NA1999.DAT.\*  
22: /SID52/gcgdata/genseq/genseqn/NA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2191	74.1	8309	18	AAV093700	Porcine retrovirus
2	2089	70.7	8196	18	AAV093699	Porcine retrovirus
3	1872	63.3	1974	18	AAV093703	Porcine retrovirus
4	1743	59.0	3482	20	AAV827349	Pig endogenous ret
5	813	27.5	7333	22	AAV77726	Defective retrovir
6	457	15.5	2462	20	AAV82748	Pig endogenous ret
7	422	14.3	7393	18	AAV74883	Porcine retrovirus
8	143	4.8	3320	18	AAV093688	Porcine retrovirus
9	98	3.3	8132	22	AAV77727	Porcine retrovirus
10	91	3.1	633	22	AAV67007	PERV-MSL LTR SEQ 1
11	91	3.1	704	22	AAV67006	IAmal LTR SEQ ID N

12	91	3.1	4918	22	AAC67022	PERV env protein c
13	91	3.1	6076	22	AAC67021	PERV env protein c
14	91	3.1	6076	22	AAC67032	PERV env protein c
15	91	3.1	7362	22	AAC67013	PERV env protein c
16	91	3.1	7873	22	AAC67023	PERV env protein c
17	91	3.1	8060	18	AAT74811	PERV env protein c
18	91	3.1	8060	22	AAE77725	Porcine retrovirus
19	80	2.7	7892	18	AAT74884	Tsukuba-1 CDNA, F
20	74	2.5	4402	22	AAC75820	Miniature swine re
21	31	1.0	2702	19	AAV55872	PERV env protein c
22	30	1.0	1431	21	AAC77334	Plant aciclovirate
23	30	1.0	1980	22	AAC677018	Human ORF328859
24	29	1.0	3557	7	AAAN60801	PERV-1.15 env prot
25	28	0.9	756	22	AA521329	Human pro-growth h
26	28	0.9	925	10	AAAN3196	Human CDNA sequen
27	28	0.9	1493	21	AAZ51086	Monkey recombinant
28	28	0.9	1736	21	AAZ61203	3'DNA of porcine e
29	28	0.9	2000	21	AAZ61087	CDNA encoding a hu
30	28	0.9	10682	21	AAZ65550	Contiguous DNA of
31	27	0.9	327	21	AAA31562	Human kinesin-like
32	27	0.9	540	22	AAH26553	plant microsatellit
33	27	0.9	579	21	AAAS0898	pig endogenous ret
34	27	0.9	599	22	AAH26255	Wheat P1 protein m
35	27	0.9	599	22	AAH26256	pig endogenous ret
36	27	0.9	599	22	AAH26258	pig endogenous ret
37	27	0.9	599	22	AAH26259	pig endogenous ret
38	27	0.9	599	22	AAH26260	pig endogenous ret
39	27	0.9	600	22	AAH26251	pig endogenous ret
40	27	0.9	600	22	AAH26252	pig endogenous ret
41	27	0.9	600	22	AAH26253	pig endogenous ret
42	27	0.9	600	22	AAH26257	pig endogenous ret
43	27	0.9	682	16	AAT02650	Met-des(F1-P2-T3-I
44	27	0.9	682	16	AAT11228	Met-des(F1-P2-T3-I
45	27	0.9	700	16	AAT02619	Met-des(Phel-Pro2

	ALIGNMENTS
RESULT 1	
AAV09700	
ID	AAV09700 standard; DNA; 8209 BP.
XX	
AC	AAV09700;
XX	
DT	20-MAY-1998 (first entry)
XX	
DE	Porcine retrovirus DNA encoding the LTR and GAG
XX	
KW	Porcine retrovirus; pREV; POL protein; ENV prot
KM	Viron core polypeptide; polymerase protein; en
XX	vaccine; diagnosis; xenotransplantation; proph
XX	
OS	Porcine retrovirus.
XX	
EH	Key
FT	Location/Qualifiers
FT	1..61
FT	/*tag= a
FT	/note= "R LTR domain"
FT	1..588
FT	/*tag= b
FT	62..143
FT	/*tag= c
FT	/note= "05 LTR domain"
FT	588..2162
FT	/*tag= d
FT	/product= GAG protein
FT	/note= "Viral core polypeptide"
FT	2163..5747
FT	/*tag= e
FT	/product= POL protein
FT	/note= "polymerase peptide"
FT	5620..7590
CD5	

```

FT      /tag= f
FT      /product= ENV protein
FT      /note= "envelope protein"
FT      LTR
FT      7591..8209
FT      /tag= g
FT      7638..8106
FT      /tag= h
FT      /note= "u3 LTR domain"
FT      7643..7686
FT      /tag= i
FT      /note= "Binding site for transcription factors
FT      OCR-1, c-Myb, Lyf-1, E47 and ETS-1"
FT      7719..7739
FT      /tag= j
FT      /note= "Binding site for transcription factors
FT      Ap-4 and NF1"
FT      7744..7754
FT      /tag= k
FT      /note= "Binding site for transcription factor AP-1/TR"
FT      7773..7785
FT      /tag= l
FT      /note= "Binding site for transcription factor ETS-1/GATA"
FT      7793..7834
FT      /tag= m
FT      /note= "Binding site for transcription factors
FT      GATA, ETS-1, c-Myb, Ap-1, GATA, E47"
FT      7841..7851
FT      /tag= n
FT      /note= "Binding site for transcription factor AP-1"
FT      7873..7883
FT      /tag= o
FT      /note= "Binding site for transcription factor AP-1"
FT      7885..7914
FT      /tag= p
FT      /note= "Binding site for transcription factors
FT      7928..7953
FT      /tag= q
FT      /note= "Binding site for transcription factors E47
FT      7964..7968
FT      /tag= r
FT      7985..7997
FT      /tag= s
FT      /note= "Binding site for transcription factor GATA"
FT      8003..8007
FT      /tag= t
FT      8024..8036
FT      /tag= u
FT      /note= "Binding site for transcription factor GATA"
FT      8042..8046
FT      /tag= v
FT      8049..8060
FT      /tag= w
FT      /note= "Binding site for transcription factor
FT      Ap-1/CRMB"
FT      8075..8090
FT      /tag= x
FT      1..61
FT      /tag= y
FT      /note= "R LTR domain"
FT      8127..8135
FT      /tag= z
FT      /note= "PolyA downstream element"
FT      8161..8166
FT      /tag= aa
XX      WO9740167-A1.
XX      30-OCR-1997.
XX      18-APR-1997;
XX      97WO-GB01087.

```

```

PR      10-FEB-1997; 97GB-0002668.
PR      19-APR-1996; 96GB-0008164.
XX      (IMUT-) IMUTRAN LTD.
PA      (QONE-) Q-ONE BIOTECH LTD.
XX      Galbraith DN, Haworth C, Lees GM, Smith KT;
XX      WPI: 1997-535851/49.
DR      P-PSDB: AAM39271; AAM39272; AAM39273.
PT      Polynucleotide encoding porcine retrovirus expression product -
PT      useful to develop products for use in vaccines, diagnosis and
PT      xeno-transplantation
XX      Claim 4; Fig 3; 69pp; English.
XX      This DNA sequence encodes the porcine retrovirus (PoEV) virion core
XX      polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and
XX      also includes the long terminal repeat (LTR). These proteins can be used
XX      to develop viral vaccines, antisense nucleic acids, ribozymes and other
XX      antiviral agents. They can also be used in xeno-transplantation
XX      technology and as diagnostic tools.
SQ      Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other:

Query Match          74.1%; Score 2191; DB 18; Length 8209;
Best local similarity 99.7%; Pred. No. 0;
Matches 2741; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY      208 gagactgcaagttccacatcgcttccaagttgagatcagttatgtatgaagccacc 267
DB      5464 gagacttgcaagttccacatcgcttccaagttgagatcagttatgtatgaagccacc 5523
QY      268 gtgcaggaaactcgaagctcggtggaaggaaccttactgtactttgaccacacaa 327
DB      5524 gtgcaggaaactcgaagctcggtggaaggaaccttactgtactttgaccacacaa 5583
QY      328 cggctgtaaaagtcgaagaacttccacctggtatcatatgcatccacgttaagctggcgc 387
DB      5584 cggctgtaaaagtcgaagaacttccacctggtatcatatgcatccacgttaagctggcgc 5643
QY      388 caactccgactcgggtgtgagagcgcgaagaagacttgaaatcccttaagtttcctccc 447
DB      5644 caactccgactcgggtgtgagagcgcgaagaagacttgaaatcccttaagtttcctccc 5703
QY      448 atcgctgttcttacttactaaacaaataactccacagcagtagtaaacgcttatagac 507
DB      5704 atcgctgttcttacttactaaacaaataactccacagcagtagtaaacgcttatagac 5763
QY      508 agctcgaaaccccatagacacttataccctacactggtcgtatattgacctgatacgggt 567
DB      5764 agctcgaaaccccatagacacttataccctacactggtcgtatattgacctgatacgggt 5823
QY      568 gtcactgttaataagaccgcgaggtgtgtccttaaggaagcactggtgacgactgcat 627
DB      5824 gtcactgttaataagaccgcgaggtgtgtccttaaggaagcactggtgacgactgcat 5883
QY      628 tctcgctcgaattatatacccgctgtttaaagacacactcccaactagtcctagtc 687
DB      5884 tctcgctcgaattatatacccgctgtttaaagacacactcccaactagtcctagtc 5943
QY      688 tatgggttctatgtcgtccagcagcagaagaagaatactgtggggttctggggaa 747
DB      5944 tatgggttctatgtcgtccagcagcagaagaagaatactgtggggttctggggaa 6003
QY      748 tctctctgtagagatgtagctgctgcctccacagatggagactggaaatgctcgatc 807
DB      6004 tctctctgtagagatgtagctgctgcctccacagatggagactggaaatgctcgatc 6063
QY      808 tctctcagaccggtaaaattctccttctgtcaattccgagcccggaagtacaaaatg 867
DB      808 tctctcagaccggtaaaattctccttctgtcaattccgagcccggaagtacaaaatg 867

```

Dh 6064 tctctccaggaacggggttaaaattctccttctgtcaattccgccccggggaagtaacaaatg 6123  
Qy 868 atgaactataataagaataagctgtccccaatccagacttagatttctaaagataagt 927  
Dh 6124 atgaactataataagaataagctgtccccaatccagacttagatttctaaagataagt 6183  
Qy 928 ttcaatgaaaaaagaacaaacagaaatatccaagaatggaataatgtaatgtaagctggga 987  
Dh 6184 ttccactg - aagaagaaacagaaatatccaagaatggaataatgtaatgtaagctggga 6241  
Qy 988 ataatltttaaataatgacgggggagcgaggttccactttaaccattcgcttaagata 1047  
Dh 6242 ataatlttltt - atataatggcgggggagcgaggttccactttaaccattcgcttaagata 6300  
Qy 1048 gagcggggagacagacccccctgtgacataggaacccataaagtaactggctgaacgggg 1107  
Dh 6301 gagcggggagacagacccccctgtgacataggaacccataaagtaactggctgaacgggg 6360  
Qy 1108 cccccggccttgagccacccagataactgacgggtgcgggtcccaatlaacctgcgtcgagct 1167  
Dh 6361 cccccggccttgagccacccagataactgacgggtgcgggtcccaatlaacctgcgtcgagct 6420  
Qy 1168 gacataacacagccgcttaacacacagtaacacactgagattgattccttaacacacgctaga 1227  
Dh 6421 gacataacacagccgcttaacacacagtaacacactgagattgattccttaacacacgctaga 6480  
Qy 1228 aactcccaagggttctcctgttaagaacaggaacagagacttccagcttcacatccagggagct 1287  
Dh 6481 aactcccaagggttctcctgttaagaacaggaacagagacttccagcttcacatccagggagct 6540  
Qy 1288 ttccaagccatacaactccacgacccctgatactgacacttctctgtgtgcttctgtatcc 1347  
Dh 6541 ttccaagccatacaactccacgacccctgatactgacacttctctgtgtgcttctgtatcc 6600  
Qy 1348 tcagggcctccttattatagtgaggagatggtcctaagaagaanaattcaatgtgaccaaagg 1407  
Dh 6601 tcagggcctccttattatagtgaggagatggtcctaagaagaanaattcaatgtgaccaaagg 6660  
Qy 1408 catagaataatcatgtaacatggggtcccggaataagcttaacccctcaactgaagtcttcggg 1467  
Dh 6661 catagaataatcatgtaacatggggtcccggaataagcttaacccctcaactgaagtcttcggg 6720  
Qy 1468 aaggagacatgcaatgaagaaagctcccccatacccaacacacctgtgctatagtaactgtg 1527  
Dh 6721 aaggagacatgcaatgaagaaagctcccccatacccaacacacctgtgctatagtaactgtg 6780  
Qy 1528 gttatagagacgctccacagaataatcgtattatgtaacctggtgataacaggtgtgggca 1587  
Dh 6781 gttatagagacgctccacagaataatcgtattatgtaacctggtgataacaggtgtgggca 6840  
Qy 1588 tgcataactggtttaaaccctgtgttccacactcagcttcaacacacacccaaagatttg 1647  
Dh 6841 tgcataactggtttaaaccctgtgttccacactcagcttcaacacacacccaaagatttg 6900  
Qy 1648 tgtgtcatggttccaacatcgcccccgagtgtaactacatccctgaggaaggtgttccttgat 1707  
Dh 6901 tgtgtcatggttccaacatcgcccccgagtgtaactacatccctgaggaaggtgttccttgat 6960  
Qy 1708 gaataagactatcgtgataacagccaaaaaagaaacccgtaacccttaaccctagctgta 1767  
Dh 6961 gaataagactatcgtgataacagccaaaaaagaaacccgtaacccttaaccctagctgta 7020  
Qy 1768 atgctcgattaggaacgccccgttggcgtfagaagaacaggaacgctgccccctgatacaga 1827  
Dh 7021 atgctcgattaggaacgccccgttggcgtfagaagaacaggaacgctgccccctgatacaga 7080  
Qy 1828 ccacagcagctagagaagaagacttgggtgactacatgacggaacatgacagaagaatctcga 1887  
Dh 7081 ccacagcagctagagaagaagacttgggtgactacatgacggaacatgacagaagaatctcga 7140  
Qy 1888 gacctaaaggagctgttggcaacctagaagaagctccctgaactcttctgtctgaagtgtt 1947  
Dh 7141 gacctaaaggagctgttggcaacctagaagaagctccctgaactcttctgtctgaagtgtt 7200

Qy 1948 ctacagaacggagggattagatctgctgttcttaagaagaagtggttattgtcaacc 2007  
Dh 7201 ctacagaacggagggattagatctgctgttcttaagaagaagtggttattgtcaacc 7260  
Qy 2008 ttaaaaagaagaatgtgtcttctatgtlagaatcaactaaggaacatacagaactccaatgaac 2067  
Dh 7261 ttaaaaagaagaatgtgtcttctatgtlagaatcaactaaggaacatacagaactccaatgaac 7320  
Qy 2068 aagcttagaanaaagttaagaagggcgtcgaagaaggaaggaagcgtgacccaaggtgtt 2127  
Dh 7321 aagcttagaanaaagttaagaagggcgtcgaagaaggaaggaagcgtgacccaaggtgtt 7380  
Qy 2128 gaagaatgttcaacaggtctcctctggaatgacacacccctctctgctctgacggggccc 2187  
Dh 7381 gaagaatgttcaacaggtctcctctggaatgacacacccctctctgctctgacggggccc 7440  
Qy 2188 ctatgtagtccctgctctgttacttaacatggtggccttgcttaataatagttgtgccc 2247  
Dh 7441 ctatgtagtccctgctctgttacttaacatggtggccttgcttaataatagttgtgccc 7500  
Qy 2248 ttgttaggaacgggtgagtgagtgacgtccagatcatalggttaaggcaacagtaacagc 2307  
Dh 7501 ttgttaggaacgggtgagtgagtgacgtccagatcatalggttaaggcaacagtaacagc 7560  
Qy 2308 ctcttgagccaaagggaacatgacactcagccttccagcttccagtttaagattagattat 2367  
Dh 7561 ctcttgagccaaagggaacatgacactcagccttccagcttccagtttaagattagattat 7620  
Qy 2368 caagaacagaagtgagggaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 2427  
Dh 7621 caagaacagaagtgagggaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 7680  
Qy 2428 aagttataataaagctctaaatgcccccggaattaaagacccctgcgtgcgtcgaat 2487  
Dh 7681 aagttataataaagctctaaatgcccccggaattaaagacccctgcgtgcgtcgaat 7740  
Qy 2488 ggtagaaggtcacacttccattctgttccagggccgtgcattcccttgcccttaagttaaac 2547  
Dh 7741 ggtagaaggtcacacttccattctgttccagggccgtgcattcccttgcccttaagttaaac 7800  
Qy 2548 aggaatagatgtgactaactgcttactgtgacttctgtaaaactgactgacacatagaag 2607  
Dh 7801 aggaatagatgtgactaactgcttactgtgacttctgtaaaactgactgacacatagaag 7860  
Qy 2608 aattgattacacatgtgacagccctagtgacctatctcaactgacatctgtcaactctgcc 2667  
Dh 7861 aattgattacacatgtgacagccctagtgacctatctcaactgacatctgtcaactctgcc 7920  
Qy 2668 aggaagccacagcagatgagcagctccggagcctatttaaaatgattgtgtccaggaagc 2727  
Dh 7921 aggaagccacagcagatgagcagctccggagcctatttaaaatgattgtgtccaggaagc 7980  
Qy 2728 gggcctcagatatttaaaatgattgtgcatgagcggcgccgtccgatattttaaat 2787  
Dh 7981 gggcctcagatatttaaaatgattgtgcatgagcggcgccgtccgatattttaaat 8040  
Qy 2788 gattgtgttgtgacgacagcgttgtgtgaaaccccaataaagctgtcccgattccgca 2847  
Dh 8041 gattgtgttgtgacgacagcgttgtgtgaaaccccaataaagctgtgtcccgattccgca 8100  
Qy 2848 ctgagggccagcagcttcttaccctgctggtgtgtgtagagcttgggccccacgagcttgg 2907  
Dh 8101 ctgagggccagcagcttcttaccctgctggtgtgtgtagagcttgggccccacgagcttgg 8160  
Qy 2908 aataaaaaatccctctgtctgtgtgcatcaaaaaaataaaaaaataa 2956  
Dh 8161 aataaaaaatccctctgtctgt 8209

RESULT 2  
AAV09699  
ID AAV09699 standard; DNA; 8196 BP.

XX AAV09699;  
AC 19-MAY-1998 (first entry)  
XX  
XX Porcine retrovirus DNA encoding, GAG, POL and ENV.  
DE Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;  
KM Porcine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
XX Porcine retrovirus.  
OS  
XX  
XX Key Location/Qualifiers  
FH 576..2126  
FT CDS /tag= a  
FT /product= GAG protein  
FT /note="viral core polypeptide"  
FT 213..5733  
FT CDS /tag= b  
FT /product= POL protein  
FT /note="polymerase peptide sequence as given in  
FT specification"  
FT 5606..7576  
FT CDS /tag= c  
FT /product= ENV protein  
FT /note="envelope protein"  
XX  
XX MO9740167-A1.  
XX  
XX 30-OCT-1997.  
XX  
XX 18-APR-1997; 97WO-GB01087.  
XX  
XX 10-FEB-1997; 97GB-0002668.  
XX 19-APR-1996; 96GB-0008164.  
XX  
XX (IMUT-) IMUTRAN LTD.  
XX (OONE-) O-ONE BIOTECH LTD.  
XX  
XX Galbraith DN, Haworth C, Lees GM, Smith KT;  
XX WPI; 1997-535851/49.  
XX  
XX Polynucleotide encoding porcine retrovirus expression product -  
PT useful to develop products for use in vaccines, diagnosis and  
PT xeno-transplantation  
XX  
PS Claim 4; Fig 2; 69pp; English.  
XX  
XX This DNA sequence encodes the porcine retrovirus (PoEV) virion core  
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These  
CC proteins can be used to develop viral vaccines, antisense nucleic acids,  
CC ribozymes and other antiviral agents. They can also be used in  
CC xeno-transplantation technology and as diagnostic tools.  
XX  
XX Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;  
SO

Query Match 70.7%; Score 2089; DB 18; Length 8196;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2739; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 208 gagactgcaagttccacatcgcttccaaagtggagattcagctatgtagagcgaac 267  
DB 5450 gagactgcaagttccacatcgcttccaaagtggagattcagctatgtagagcgaac 267  
QY 268 gtgcaggaacactcgagctcggtggaagggaaccttctctgtaatttgaccaccaa 327  
DB 5510 gtgcaggaacactcgagctcggtggaagggaaccttctctgtaatttgaccaccaa 327  
QY 328 cggctgtgaagtggaaggaatccacactgagatcagatgacacgaagtgagcgc 387  
DB 5570. cggctgtgaagtggaaggaatccacactgagatcagatgacacgaagtgagcgc 387

QY 388 caactccgactcgggtgtagagcgaagaaagactggaatcccttaagcttcgctcc 447  
DB 5630 caactccgactcgggtgtagagcgaagaaagactggaatcccttaagcttcgctcc 5689  
QY 448 atcgctgttccctactcttaacaaatacctccagccagtagtaaacgctctagac 507  
DB 5690 atcgctgttccctactcttaacaaatacctccagccagtagtaaacgctctagac 5749  
QY 508 agctggaaccccatagaccttataccctactgctgattatgacccgatacgggt 567  
DB 5750 agctggaaccccatagaccttataccctactgctgattatgacccgatacgggt 5809  
QY 568 gtcaactgataatagcactcgaagtggtgtccctagaggaacctgtgtgacctgacat 627  
DB 5810 gtcaactgataatagcactcgaagtggtgtccctagaggaacctgtgtgacctgacat 5869  
QY 628 ttctgctccgattgataccccgcgtgttaaaagacacccctccaaactgctcgtagt 687  
DB 5870 ttctgctccgattgataccccgcgtgttaaaagacacccctccaaactgctcgtagt 5929  
QY 688 tatgggttctatgtgtcccaaggcacaagaaagaaatactgtggtggtctgaggaa 747  
DB 5930 tatgggttctatgtgtcccaaggcacaagaaagaaatactgtggtggtctgaggaa 5989  
QY 748 tccctctgtagagatgtagagctgctgacccccaagatgtagactggaatgagcagatc 807  
DB 5990 tccctctgtagagatgtagagctgctgacccccaagatgtagactggaatgagcagatc 6049  
QY 808 tctctcgaagacccgggttaaattctctctgtgtcaatttcggccggcggaagtagaagt 867  
DB 6050 tctctcgaagacccgggttaaattctctctgtgtcaatttcggccggcggaagtagaagt 6109  
QY 868 atgaaactatataagataagagctgctccccaagactttagattatataagaatagt 927  
DB 6110 atgaaactatataagataagagctgctccccaagactttagattatataagaatagt 6169  
QY 928 ttcaactgaaagaaagaaacagaaataatcctaaagtgtgataatgtgtgctgaggaa 987  
DB 6170 ttcaactg--aaagaaagaaagaaataatcctaaagtgtgataatgtgtgctgaggaa 6227  
QY 988 atagttttttaaataatgtagcgggagagaggttccaaacttcaacttgccttggagata 1047  
DB 6228 atagttttt-attataatgtagcgggagagaggttccaaacttcaacttgccttggagata 6286  
QY 1048 gagacgggagacagaacccctgttggaatgtagaccggaataaagtactggtgtaacagggt 1107  
DB 6287 gagacgggagacagaacccctgttggaatgtagaccggaataaagtactggtgtaacagggt 6346  
QY 1108 ccccgccctgtagcagccacgataattgctgcgggtggcccaatttaacctcgttggcgt 1167  
DB 6347 ccccgccctgtagcagccacgataattgctgcgggtggcccaatttaacctcgttggcgt 6406  
QY 1168 gacataacacagccgcttagcacaacgttaccctgtgattatccttaccacaacgcttaga 1227  
DB 6407 gacataacacagccgcttagcacaacgttaccctgtgattatccttaccacaacgcttaga 6466  
QY 1228 aactcccaagtggtctcgttlaagacagagacagacttctcaagtcataccaggaagt 1287  
DB 6467 aactcccaagtggtctcgttlaagacagagacagacttctcaagtcataccaggaagt 6526  
QY 1288 ttccaagcacaactcacaacgacccctgtgccaattctctgtgtgtgtgtgtatcc 1347  
DB 6527 ttccaagcacaactcacaacgacccctgtgccaattctctgtgtgtgtgtgtatcc 6586  
QY 1348 tcaagggtctcctattatagaggtgtagtgaagaaagaaataatcaatttgacaaagag 1407  
DB 6587 tcaagggtctcctattatagaggtgtagtgaagaaagaaataatcaatttgacaaagag 6646  
QY 1408 catagaatcaatgatacaggggtgtccgaaataaacttaccctcaactgaagtgtccggg 1467  
DB 6647 catagaatcaatgatacaggggtgtccgaaataaacttaccctcaactgaagtgtccggg 6706





CC isolated from the human cell line Raji. Such viral proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.

XX Sequence 1974 BP: 534 A; 489 C; 483 G; 468 T; 0 other;

Query Match 63.3%; Score 1872; DB 18; Length 1974;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgacccacgcttaagctggcgccaccccgacccgctgggggtgagagacggaaaagactg 423  
 Db 1 atgacccacgcttaagctggcgccaccccgacccgctgggggtgagagacggaaaagactg 60  
 QY 424 agaattcccttaagcttcgcctcctacgtcgtgttccttacttaacaataactcccgag 483  
 Db 61 agaattcccttaagcttcgcctcctacgtcgtgttccttacttaacaataactcccgag 120  
 QY 484 gccagtagtaacgacctatagacagctcgaaaccccatagacctttatcccttaacttg 543  
 Db 121 gccagtagtaacgacctatagacagctcgaaaccccatagacctttatcccttaacttg 180  
 QY 544 ctgattatgacccctgtagcgggtgtcactgttaaatagcactcgagggtgtgtcctaga 603  
 Db 181 ctgattatgacccctgtagcgggtgtcactgttaaatagcactcgagggtgtgtcctaga 240  
 QY 604 ggcacccgtgtgctgtaacgtgatttctgcctccgattgattaaacccgctgttaaaagc 663  
 Db 241 ggcacccgtgtgctgtaacgtgatttctgcctccgattgattaaacccgctgttaaaagc 300  
 QY 664 acaccccccacacccagtcgagtagtctggtctcattgctgcgcccagggcacagagaagag 723  
 Db 301 acaccccccacacccagtcgagtagtctggtctcattgctgcgcccagggcacagagaagag 360  
 QY 724 aaatactgtgggggtctctggggaatcctctctgtagagatgagctgtgcctaccac 783  
 Db 361 aaatactgtgggggtctctggggaatcctctctgtagagatgagctgtgcctaccac 420  
 QY 784 gatggagactggaatgagccgactctctccacagacccgggttaaatcttctctgttcaat 843  
 Db 421 gatggagactggaatgagccgactctctccacagacccgggttaaatcttctctgttcaat 480  
 QY 844 tccggcccgagcaagtaacaaatgataaataataaagaagctgtcccccata 903  
 Db 481 tccggcccgagcaagtaacaaatgataaataataaagaagctgtcccccata 540  
 QY 904 gacttagattactaaagataagtttcaactgaaagaaagaaacagagaaatattcaaaag 963  
 Db 541 gacttagattactaaagataagtttcaactgaaagaaagaaacagagaaatattcaaaag 600  
 QY 964 tggataaatgtagtagctgggggaatagtttttaataatattggtggggagacgggtcc 1023  
 Db 601 tggataaatgtagtagctgggggaatagtttttaataatattggtggggagacgggtcc 660  
 QY 1024 acttaaacactgccttagatagagacgggacgaacccctgtggaatggagacc 1083  
 Db 661 acttaaacactgccttagatagagacgggacgaacccctgtggaatggagacc 720  
 QY 1084 gataaagtactgtgtgacaggggccccggccctgtgagccacgcgcaataactgtgcggtg 1143  
 Db 721 gataaagtactgtgtgacaggggccccggccctgtgagccacgcgcaataactgtgcggtg 780  
 QY 1144 ccccaatlaacctgctggtgctgacataacacagccgcttagcaacagtagtaacctgga 1203  
 Db 781 ccccaatlaacctgctggtgctgacataacacagccgcttagcaacagtagtaacctgga 840  
 QY 1204 ttgattctactaacacagccttagaagaactcccaaggtgttctgtttaagaagagacagaga 1263  
 Db 841 ttgattctactaacacagccttagaagaactcccaaggtgttctgtttaagaagagacagaga 900  
 QY 1264 ctcttcagctcatccagggagctttccaagccaatccaactccagccctgtatgacct 1323

Db 901 ctcttcagctcatccagggagctttccaagccaatccaactccagccctgtatgacct 960  
 QY 1324 tctcttgtgtgctgtgtctatccacagggcctctcttattatgagggatgctaaagaa 1383  
 Db 961 tctcttgtgtgctgtgtctatccacagggcctctcttattatgagggatgctaaagaa 1020  
 QY 1384 agaaaaatcaatgtgacccaagagacataaataatgatactgtgggggtccccaataag 1443  
 Db 1021 agaaaaatcaatgtgacccaagagacataaataatgatactgtgggggtccccaataag 1080  
 QY 1444 cttaacctactgaaagttccgggaaggagacatgacataggaagaagctccccaatccac 1503  
 Db 1081 cttaacctactgaaagttccgggaaggagacatgacataggaagaagctccccaatccac 1140  
 QY 1504 caaaccttgcctatgactgtgtgtttatgagcagggcctcagaagaatcagatattagta 1563  
 Db 1141 caaaccttgcctatgactgtgtgtttatgagcagggcctcagaagaatcagatattagta 1200  
 QY 1564 cctgtgtataacagaggtgtggcaltgcaataactgtgggtlaacccctgtgttccacctca 1623  
 Db 1201 cctgtgtataacagaggtgtggcaltgcaataactgtgggtlaacccctgtgttccacctca 1260  
 QY 1624 gtcttcaacaaatccaagaattgtgtgtcatgtgtccaaatcgtcccccagtgtaactac 1683  
 Db 1261 gtcttcaacaaatccaagaattgtgtgtcatgtgtccaaatcgtcccccagtgtaactac 1320  
 QY 1684 catcctgagaagatgtgtctcttgatgataatgactctcgttataacggccaaagaagaa 1743  
 Db 1321 catcctgagaagatgtgtctcttgatgataatgactctcgttataacggccaaagaagaa 1380  
 QY 1744 cccgtatcccttaaccccttaactgttaatgtccgagatagagagacggctgtgcaataagaa 1803  
 Db 1381 cccgtatcccttaaccccttaactgttaatgtccgagatagagagacggctgtgcaataagaa 1440  
 QY 1804 gggacagctgcccctgataccaaggaaccaagcagctagagaagaagactgtgtgactacat 1863  
 Db 1441 gggacagctgcccctgataccaaggaaccaagcagctagagaagaagactgtgtgactacat 1500  
 QY 1864 gggcccatagacgaagaatcccgagccttaagaagggtcgtgttagaaccctagaagaagtc 1923  
 Db 1501 gggcccatagacgaagaatcccgagccttaagaagggtcgtgttagaaccctagaagaagtc 1560  
 QY 1924 ctgactctctgtctgaagtggttctacagaacccggaggggattagatctgtcttcta 1983  
 Db 1561 ctgactctctgtctgaagtggttctacagaacccggaggggattagatctgtcttcta 1620  
 QY 1984 agagaagtggtgtatgtgacgcttaagaagaatgtgtcttctatgtagatactca 2043  
 Db 1621 agagaagtggtgtatgtgacgcttaagaagaatgtgtcttctatgtagatactca 1680  
 QY 2044 ggaagccatagagactccatgacaaagcttaaaaaaagttagaaggcgttgaaggaa 2103  
 Db 1681 ggaagccatagagactccatgacaaagcttaaaaaaagttagaaggcgttgaaggaa 1740  
 QY 2104 agagaagctgacaaaggggtgttgaagatgttccaagaagcttccctgtgatacgaacc 2163  
 Db 1741 agagaagctgacaaaggggtgttgaagatgttccaagaagcttccctgtgatacgaacc 1800  
 QY 2164 ctgcttctgtcctgacaggggccccctagtagctcgtctcgttacttaacagtgtggcct 2223  
 Db 1801 ctgcttctgtcctgacaggggccccctagtagctcgtctcgttacttaacagtgtggcct 1860  
 QY 2224 tgccttaataatagattgtgttccttcttgaagaagagatgagtagcagctcaagatag 2283  
 Db 1861 tgccttaataatagattgtgttccttcttgaagaagagatgagtagcagctcaagatag 1920  
 QY 2284 gtacttagcaacagtagcaagaagccttctgacaaagagagaacagactgacctag 2337  
 Db 1921 gtacttagcaacagtagcaagaagccttctgacaaagagagaacagactgacctag 1974

RESULT 4

AAV82749  
 ID AAV82749 standard; DNA; 3482 BP.  
 AC AAV82749;  
 XX 25-FEB-1999 (first entry)  
 XX  
 XX Pig endogenous retrovirus (PERV)-B envelope (env) gene region.  
 DE  
 XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
 KW probe; detection; retrovirus; human tissue; xenotransplant;  
 KW primary porcine tissue; human cell line; porcine cell line; ss.  
 XX  
 OS Pig endogenous retrovirus.  
 XX  
 FH Key Location/Qualifiers  
 FT 911.2884 /\*tag= a  
 FT CDS /product= envelope\_protein  
 XX  
 XX WO9853104-A2.  
 XX  
 XX 26-NOV-1998.  
 XX  
 XX 18-MAY-1998; 98WO-GB01428.  
 XX  
 XX 16-MAY-1997; 97GB-0010154.  
 XX  
 XX (MED1-) MEDICAL RES COUNCIL.  
 XX  
 XX Stoye JP, Weiss RA;  
 XX  
 XX MPI; 1999-045324/04.  
 DR P-PSDB; AAW85452.  
 XX  
 XX Newly isolated nucleic acid probe capable of hybridising to either  
 PT the PERV-A or PERV-B env gene - useful in the detection of  
 PT retroviruses, and their subtypes, in a sample of porcine/human  
 PT tissue  
 XX  
 XX Claim 3; Page 21-23; 36pp; English.  
 PS  
 XX The present sequence encodes a pig endogenous retrovirus (PERV)-B  
 CC envelope protein. PERV exists in two different subtypes, PERV-A and  
 CC PERV-B. The differences are reflected in sequence divergence in the  
 CC envelope genes. Probes and primers can be derived from the envelope  
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
 CC a method to detect retroviruses in a sample of porcine/human tissue,  
 CC particularly primary porcine tissue and human cell lines that have been  
 CC cultivated in the presence of a porcine cell line, or human tissue from  
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
 CC one of the PERV env genes can also be determined.  
 CC  
 CC Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;  
 SQ

Query Match 59.0%; Score 1743; DB 20; Length 3482;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

Db	728	agctccggagggcctactcagaggaaggagacttgcaagttccacatgcttccaaagtgg	787
Qy	241	gagattcagctctatgttagacgcccacgltgcaagaaacctcgagactcgttggaaaggac	300
Db	788	gagattcagctctatgttagacgcccacgltgcaagaaacctcgagactcgttggaaaggac	847
Qy	301	cttatctgacttttgaccacacaaagcgtgtgtgaagtgcgaagaaatccacactgga	360
Db	848	cttatctcgtactttgaccacacaaagcgtgtgtgaagtgcgaagaaatccacactgga	907
Qy	361	ttcatgcatccacglttaagctgtgcgcaacctccgacccggaggtggagagccgaaaga	420
Db	908	ttcatgcatccacglttaagctgtgcgcaacctccgacccggaggtggagagccgaaaga	967
Qy	421	ctgagaatcccttaagacttcgctccatcgcctgtgttcttacttaacataactccc	480
Db	968	ctgagaatcccttaagacttcgctccatcgcctgtgttcttacttaacataactccc	1027
Qy	481	caaggcagtagtaaaagcctttagacagctcgaacccccatagaccttaccctacc	540
Db	1028	caaggcagtagtaaaagcctttagacagctcgaacccccatagaccttaccctacc	1087
Qy	541	tggtgatattgacccctgatalacggtgtcactgttaaatagcactcgaagtgtgtcct	600
Db	1088	tggtgatattgacccctgatalacggtgtcactgttaaatagcactcgaagtgtgtcct	1147
Qy	601	agaggaacctgtgtgacctgaactgcattctgcctccgattgataaaccgcgtttaa	660
Db	1148	agaggaacctgtgtgacctgaactgcattctgcctccgattgataaaccgcgtttaa	1207
Qy	661	agacacacctcccaacctagctcgttagttgtgtctatgtctgcgccggacaaagaa	720
Db	1208	agacacacctcccaacctagctcgttagttgtgtctatgtctgcgccggacaaagaa	1267
Qy	721	gagaataactgtgggggttcctgggaatccctctgtgagagatggagctgcgtaccctc	780
Db	1268	gagaataactgtgggggttcctgggaatccctctgtgagagatggagctgcgtaccctc	1327
Qy	781	aacgatgagagactggaatlgccgatctctccagagaccgggtaaatctctctgtc	840
Db	1328	aacgatgagagactggaatlgccgatctctccagagaccgggtaaatctctctgtc	1387
Qy	841	aattccggccggggaagaaatgataagacatataagataagagctgcctccca	900
Db	1388	aattccggccggggaagaaatgataagacatataagataagagctgcctccca	1447
Qy	901	tcaagcttagattcctaagaatagttcactgtaaaagaagaaacagaaataattcaa	960
Db	1448	tcaagcttagattcctaagaatagttcactgtaaaagaagaaacagaaataattcaa	1507
Qy	961	aagtgtgataaattgatagtgagctggggaatagtttcttaataatcgtgcggggagcagg	1020
Db	1508	aagtgtgataaattgatagtgagctggggaatagtttcttaataatcgtgcggggagcagg	1567
Qy	1021	ttcaccttaaacattcgccttagagtagagagagggagacaaacccctgtggcaattgga	1080
Db	1568	ttcaccttaaacattcgccttagagtagagagagggagacaaacccctgtggcaattgga	1627
Qy	1081	cccgaataaagtaactgtgctgaacaagggcccccgggcccgcacacgataacttgcg	1140
Db	1628	cccgaataaagtaactgtgctgaacaagggcccccgggcccgcacacgataacttgcg	1687
Qy	1141	gtgcccaattaaactcgtcgtgcgcctgacataacacagccgactagcaacagtaacact	1200
Db	1688	gtgcccaattaaactcgtcgtgcgcctgacataacacagccgactagcaacagtaacact	1747
Qy	1201	ggattgattcctaccacacagccctagaaatccccagggttctcgtttaaagagggacag	1260
Db	1748	ggattgattcctaccacacagccctagaaatccccagggttctcgtttaaagagggacag	1807
Qy	1261	agactcttcagctcatccaggaagcttccaagccatcaactccacgacccgtgatgcc	1320



XX The present invention relates to a method for screening a cell or tissue  
 CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g. the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.

SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 27.5%; Score 813; DB 22; Length 7333;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 863; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2082 gttagaagagcgtcgaaggaagagagctaccaggggtgttgaagagtgctcaa 2141  
 Db 6470 gttagaagagcgtcgaaggaaggaagagcgcagcaggggtgttgaagagtgctcaa 6529  
 OY 2142 caggtctccttgatgacacacccctgtcttctgtctgtacggggccctagtatctgtct 2201  
 Db 6530 caggtctccttgatgacacacccctgtcttctgtctgtacggggccctagtatctgtct 6589  
 OY 2202 cctgttacttaacagttgggcttgccttaataagttgttgccttgttagaagacg 2261  
 Db 6590 cctgttacttaacagttgggcttgccttgccttaataagttgttgccttgttagaagacg 6649  
 OY 2262 agtgaatgagctccagatatagttacttaggcaacagatccagagcccttctgtacgaag 2321  
 Db 6650 agtgaatgagctccagatatagttacttaggcaacagatccagagcccttctgtacgaag 6709  
 OY 2322 agaaactgacacttagccttccagcttcttaagattagaaacttaacaagaacaagtg 2381  
 Db 6710 agaaactgacacttagccttccagcttcttaagattagaaacttaacaagaacaagtg 6769  
 OY 2382 ggggaatgaaagatgaaatgcaacctaaacctcccaagaccaggaagttaataaag 2441  
 Db 6770 ggggaatgaaagatgaaatgcaacctaaacctcccaagaccaggaagttaataaag 6829  
 OY 2442 cctaataatgcccccaataatacagacccctgtgtcgcgcgaataaagttagaagtgaca 2501  
 Db 6830 cctaataatgcccccaataatacagacccctgtgtcgcgcgaataaagttagaagtgaca 6889  
 OY 2502 cttcctatgttccagggcctgtctacctgtgcctaaagtaagaatacaggaatgagtgca 2561  
 Db 6890 cttcctatgttccagggcctgtctacctgtgcctaaagtaagaatacaggaatgagtgca 6949  
 OY 2562 ctaatgccttactctgagttctgttaaaactgactgacacatagaagaattgtatcacat 2621  
 Db 6950 ctaatgccttactctgagttctgtgttaaaactgactgacacatagaagaattgtatcacat 7009  
 OY 2622 tgaacagccctaaatgacatactcaactgaatcttgaactgtgccccaggaagccccagag 2681  
 Db 7010 tgaacagccctaaatgacatactcaactgaatcttgaactgtgccccaggaagccccagag 7069  
 OY 2682 atgcgagacctccgagactatttaaaatgatgtgtccacagagcgcggtctcgatatc 2741  
 Db 7070 atgcgagacctccgagactatttaaaatgatgtgtccacagagcgcggtctcgatatc 7129  
 OY 2742 ttaaaatgatgtgtccacagagcgcggtctcgatatatttaaaatgatgtgtgtgac 2801  
 Db 7130 ttaaaatgatgtgtccacagagcgcggtctcgatatatttaaaatgatgtgtgtgac 7189  
 OY 2802 gacagagccttgtgtgaacccataaagcgtgtccgatttcgacagcgggcccgcagcgt 2861  
 Db 7190 gacagagccttgtgtgaacccataaagcgtgtccgatttcgacagcgggcccgcagcgt 7249  
 OY 2862 ccttaacccctgtgtgtgactgtgtgagccacagcgcgttgaataaaatcctct 2921

Db 7250 ccttaacccctgtgtgtgactgtgtgagccacagcgcgttgaataaaatcctct 7309  
 OY 2922 tgcgtgtgtcatcaaaaaaaa 2945  
 Db 7310 tgcgtgtgtcatcaaaaaaaa 7333

# RESULT 6

ID AAV82748 standard; DNA; 2462 BP.

AAV82748;

25-FEB-1999 (first entry)

DE Pig endogenous retrovirus (PERV)-A envelope (env) gene region.

KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;

KW probe; primer; detection; retrovirus; human tissue; xenotransplant;

XX primary porcine tissue; human cell line; porcine cell line; ss.

OS Pig endogenous retrovirus.

XX Key Location/Qualifiers

FT CDS 211..2193

FT /tag= a

FT /product= envelope-protein

PD WO9853104-A2.

PD 26-NOV-1998.

PE 18-MAY-1998; 98WO-GB01428.

PR 16-MAY-1997; 97GB-0010154.

PA (MED- ) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPL: 1999-045324/04.

DR P-PSDB: AAM85453.

XX Newly isolated nucleic acid probe capable of hybridizing to either

PT the PERV-A or PERV-B env gene - useful in the detection of

PT retroviruses, and their subtypes, in a sample of porcine/human

PT tissue

PS Claim 6; Page 20-21; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-A

CC envelope protein. PERV exists in two different subtypes, PERV-A and

CC PERV-B. The differences are reflected in sequence divergence in the

CC envelope genes. Probes and primers can be derived from the envelope

CC (env) genes of PERV-A and PERV-B. The probes and primers are used in

CC a method to detect retroviruses in a sample of porcine/human tissue,

CC particularly primary porcine tissue and human cell lines that have

CC been cultivated in the presence of a porcine cell line, or human tissue

CC from a patient with a xenotransplant. Subtype of PERV in a sample containing

CC one of the PERV env genes can also be determined.

XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match 15.5%; Score 457; DB 20; Length 2462;

Best Local Similarity 99.8%; Pred. No. 2,6e-178;

Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2082 gttagaagagcgtcgaaggaagagagcgtcaccaggggtgttgaagagtgctcaa 2141  
 Db 1938 gttagaagagcgtcgaaggaagagagcgtcaccaggggtgttgaagagtgctcaa 1997

QY	2142	caagtcctccttgatgtagcaacccctgctcttcctcgtacgaggggccccttagtagctcgtc	2201
Db	1998	caagtcctccttgatgtagcaacccctgctcttcctcgtacgaggggccccttagtagctcgtc	2057
QY	2202	cctgttaacttaacagtttvggsccttgcttaataaagtttctgtgctctgttatagagaag	2261
Db	2058	cctgttaacttaacagtttvggsccttgcttaataaagtttctgtgctctgttatagagaag	2117
QY	2252	agttagtgcagctccaagatactatggtactatggccaacgatacccaagccttcgtagccaag	2321
Db	2118	agttagtgcagctccaagatactatggtactatggccaacgatacccaagccttcgtagccaag	2177
QY	2322	agaaacgacacctctagaccttcgccagcttccttaagaattagaactatcaagaacaagaatg	2381
Db	2178	agaaacgacacctctagaccttcgccagcttccttaagaattagaactatcaagaacaagaatg	2237
QY	2382	gggagatgaagaatgaaatagcaacctaaacctcccaagaccagaagattataaagaag	2441
Db	2238	gggagatgaagaatgaaatagcaacctaaacctcccaagaccagaagattataaagaag	2297
QY	2442	ctctaaatgcccccggaattacagacctctgctgctgcagatgaattagttaaagaagta	2501
Db	2298	ctctaaatgcccccggaattacagacctctgctgctgcagatgaattagttaaagaagta	2357
QY	2502	cttcctcatgtttccagagcctgtcatlccggcccttaagtaagaataacagaagaatgattga	2561
Db	2358	cttcctcatgtttccagagcctgtcatlccggcccttaagtaagaataacagaagaatgattga	2417
QY	2562	ctaatgccttatctcgattcttgyaaaac	2589
Db	2418	ctaatgccttatctcgattcttgyaaaac	2445

RESULT	7
AAT74883	
ID	AAT74883 standard; cDNA; 7393 BP.
XX	
AC	AAT74883;
XX	
DT	09-FEB-1998 (first entry)
XX	
DE	Porcine retrovirus cDNA (defective).
XX	
KM	Retrovirus; porcine; GAG protein; POL protein; ENV protein; oncotransplantation; infectious; provirus; organ transplant; donor activated virus; PCR; ss.
XX	
OS	Porcine retrovirus.
XX	
FH	Location/Qualifiers
FT	CDS 598..2172
FT	/tag= a
FT	mat_peptide 598..2169
FT	/tag= b
FT	/note= "putative GAG protein"
FT	CDS 2320..4737
FT	/tag= c
FT	/note= "putative POL coding region (partial) as described in the specification"
FT	mat_peptide 2320..3522
FT	/tag= d
FT	/note= "putative POL protein (partial)"
FT	mat_peptide 3516..4328
FT	/tag= e
FT	/note= "putative POL protein (partial)"
FT	mat_peptide 4332..4748
FT	/tag= f
FT	CDS 4738..6725
FT	/tag= g
FT	/note= "putative ENV coding region (partial) as described in the specification"
FT	mat_peptide 4752..6722
FT	/tag= h

FT		/note= "ENV protein (partial)"
XX		
PN	W09721836-A1.	
XX		
PD	19-JUN-1997.	
XX		
PF	13-DEC-1996;	96WO-US19680.
XX		
PR	14-DEC-1995;	95US-0572645.
XX		
PA	(GEHO ) GEN HOSPITAL CORP.	
XX		
PI	Fishman JA;	
XX		
DR	WPI; 1997-332804/30.	
DR	P-PSDB; AAW32091-W32095.	
XX		
PT	New nucleic acid from porcine retroviruses - used for detecting	
PT	viruses in transplant or other tissue and for assessing risk of	
PT	transmitting infection to graft recipient	
XX		
PS	Claim 16; Fig 2; 128pp; English.	

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
14.3%; Score 422; DB 18; Length 7393;	99.8%; Pred. No. 5,3e-164;	472;	0;	1;	0;	0;
2082	gtttagagagcgctcgaaaggaaagagagctgacaaagggtgtgttgaagatgttcaa	2141				
6470	gttttagagagcgctcgaaaggaaagagagctgacaaagggtgtgttgaagatgttcaa	6529				
2142	caggtctccttggatgacacacccttcttctgtcttgaaggggccctagtgtctgtct	2201				
6530	caggtctccttggatgacacacccttcttctgtcttgaaggggccctagtgtctgtct	6589				
2202	cctgttacttaacgtttgggccttgtttaaataatgattgtgtgttctgtttagaagac	2261				
6590	cctgttacttaacgtttgggccttgtttaaataatgattgtgtgttctgtttagaagac	6649				
2262	agtgtgtcagctccagatcatgttacttaggacaacagtcaaaagcctctgagccaag	2321				
6650	agtgtgtcagctccagatcatgttacttaggacaacagtcaaaagcctctgagccaag	6709				
2322	agaaactcagctctgacctccagttcttaagattagaactcttaacaaagacaagaatg	2381				
6710	agaaactcagctctgacctccagttcttaagattagaactcttaacaaagacaagaatg	6769				
2382	gggaatgaaagatataaaatgtaaccttaacctccagaaacccagaagattataaaaag	2441				
6770	gggaatgaaagatataaaatgtaaccttaacctccagaaacccagaagattataaaaag	6829				



RESULT 10  
AAC67007  
ID AAC67007 standard; DNA; 633 BP.  
XX  
XX  
AC AAC67007;  
XX  
XX  
DT 27-MAR-2001 (first entry)  
XX  
XX  
DE PERV-MSL LTR SEQ ID NO: 7.  
XX  
XX  
KW Xenotransplantation; infectious agent; vaccine; ds.  
XX  
OS Porcine endogenous retrovirus.  
XX  
PN WO200071726-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 24-MAY-2000; 2000WO-US14296.  
XX  
PR 24-MAY-1999; 99US-0135631.  
XX  
PA (MAYO-) MAYO MEDICAL VENTURES.  
XX  
PI Federspiel MJ;  
XX  
DR WPI; 2001-032041/04.  
XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
XX  
XX  
PS Disclosure; Page 98; 144pp; English.  
XX  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX  
XX  
SQ Sequence 633 BP; 170 A; 125 C; 147 G; 191 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 633;  
Best Local Similarity 100.0%; Pred. No. 5.7e-28;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagcagctctaccctcgctgtgtacgactgtggcccgagcgctt 2905  
|||||  
DB 452 cactcgaggcgagcagctctaccctcgctgtgtacgactgtggcccgagcgctt 511  
|||||  
QY 2906 ggaataaatacctctgtctgttgcacaa 2936  
|||||  
DB 512 ggaataaatacctctgtctgttgcacaa 542  
|||||

RESULT 11  
AAC67006  
ID AAC67006 standard; DNA; 704 BP.  
XX  
XX  
AC AAC67006;  
XX  
XX  
DT 27-MAR-2001 (first entry)  
XX  
XX  
DE 1amA1 LTR SEQ ID NO: 6.  
XX  
XX  
KM Xenotransplantation; infectious agent; vaccine; ds.  
XX  
XX  
OS Unidentified.  
XX  
PN WO200071726-A1.  
XX

PD 30-NOV-2000.  
XX  
XX  
PF 24-MAY-2000; 2000WO-US14296.  
XX  
XX  
PR 24-MAY-1999; 99US-0135631.  
XX  
XX  
PA (MAYO-) MAYO MEDICAL VENTURES.  
XX  
XX  
PI Federspiel MJ;  
XX  
DR WPI; 2001-032041/04.  
XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
XX  
XX  
PS Disclosure; Page 97-98; 144pp; English.  
XX  
XX  
CC The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX  
XX  
SQ Sequence 704 BP; 187 A; 162 C; 158 G; 197 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 704;  
Best Local Similarity 100.0%; Pred. No. 5.6e-28;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagcagctctaccctcgctgtgtacgactgtggcccgagcgctt 2905  
|||||  
DB 523 cactcgaggcgagcagctctaccctcgctgtgtacgactgtggcccgagcgctt 582  
|||||  
QY 2906 ggaataaatacctctgtctgttgcacaa 2936  
|||||  
DB 583 ggaataaatacctctgtctgttgcacaa 613  
|||||

RESULT 12  
AAC67022  
ID AAC67022 standard; DNA; 4918 BP.  
XX  
XX  
AC AAC67022;  
XX  
XX  
DT 27-MAR-2001 (first entry)  
XX  
XX  
DE PERV env protein coding sequence SEQ ID NO: 22.  
XX  
XX  
KM Xenotransplantation; infectious agent; vaccine; ds.  
XX  
XX  
OS Porcine endogenous retrovirus.  
XX  
PN WO200071726-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 24-MAY-2000; 2000WO-US14296.  
XX  
XX  
PR 24-MAY-1999; 99US-0135631.  
XX  
XX  
PA (MAYO-) MAYO MEDICAL VENTURES.  
XX  
XX  
PI Federspiel MJ;  
XX  
DR WPI; 2001-032041/04.  
XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -



XX Claim 16; Page 109-111; 144pp; English.  
PS  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX  
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 4918;  
Best Local Similarity 100.0%; Pred. No. 4.3e-28;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 2905  
DB 4736 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 4795  
QY 2906 ggaataaaaatcctctgtctgttgcataca 2936  
DB 4796 ggaataaaaatcctctgtctgttgcataca 4826

## RESULT 13

AAC67021  
ID AAC67021 standard; DNA: 6076 BP.

AC AAC67021;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 21.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

PF 24-MAY-2000; 2000WO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -

XX Claim 16; Page 107-109; 144pp; English.

CC The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX

SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 6076;  
Best Local Similarity 100.0%; Pred. No. 4.2e-28;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 2905  
DB 572 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 631  
QY 2906 ggaataaaaatcctctgtctgttgcataca 2936  
DB 632 ggaataaaaatcctctgtctgttgcataca 662

## RESULT 14

AAC67032  
ID AAC67032 standard; DNA: 6076 BP.

AC AAC67032;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 32.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

PF 24-MAY-2000; 2000WO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -

XX Claim 16; Page 117-119; 144pp; English.

CC The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX

SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 3.1%; Score 91; DB 22; Length 6076;  
Best Local Similarity 100.0%; Pred. No. 4.2e-28;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2846 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 2905  
DB 572 cactcgaggcgagctcctctacccctgctggtgtaacagactgtggcccaagcgactt 631  
QY 2906 ggaataaaaatcctctgtctgttgcataca 2936  
DB 632 ggaataaaaatcctctgtctgttgcataca 662

## RESULT 15

AAC67019  
ID AAC67019 standard; DNA: 7362 BP.

AC AAC67019;

DT 27-MAR-2001 (first entry)

XX	PERV env protein coding sequence SEQ ID NO: 19.
DE	
XX	Xenotransplantation; infectious agent; vaccine; ds.
KW	
XX	Porcine endogenous retrovirus.
OS	
XX	MO200071726-A1.
PN	
XX	
PD	30-NOV-2000.
XX	
PF	24-MAY-2000; 2000MO-US14296.
XX	
PR	24-MAY-1999; 99US-0135631.
XX	
XX	
PA	(MAYO-) MAYO MEDICAL VENTURES.
XX	
PI	Federspiel MJ;
XX	
DR	WPI: 2001-032041/04.
XX	
PT	Inhibiting or preventing infectious agent transmission in mammalian
PT	transplant recipients, by introducing recombinant DNA comprising DNA
PI	encoding extracellular proteins of the agent into donor cells, such as
PT	swine cells -
XX	
XX	
PS	Claim 16; Page 101-104; 144pp; English.
XX	
CC	The present invention provides a method to prevent the transmission of
CC	infectious agents during xenotransplantation. This involves introducing
CC	to donor swine cells a recombinant DNA encoding a peptide fragment from
CC	the infectious agent, and then introducing these cells into the
CC	transplant recipient.
XX	
XX	
SEQ	Sequence 7362 BP; 1997 A; 1821 C; 1861 G; 1663 T; 0 other;

	Query Match	3.18;	Score 91;	DB 22	Length 7362;
	Best Local Similarity	100.0%;	Pred. No. 4.1e-28;		
	Matches 91;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps
QY	2846	cactcggcgagccgagctctaccctctggtgtgtacgactgtgagccacgacgtt	2905		
Db	637	cactcggcgagccgagctctaccctctggtgtgtacgactgtgagccacgacgtt	696		
QY	2906	ggaataaaaatcctctgtcgtttgcatcaa	2936		
Db	697	ggaataaaaatcctctgtcgtttgcatcaa	727		

Search completed: February 24, 2002, 07:43:57  
Job time: 16942 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2002, 21:40:34 ; Search time 624.99 Seconds

(without alignments)  
4554.185 Million cell updates/sec

Title: US-09-171-553B-1

Perfect score: 3320

Sequence: 1 gaattcgcgccgctgcgc.....aaaaaaaaaaaaaaaaa 3320

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3320	100.0	3320	18	AAV09698 Porcine retrovirus
2	3179	95.8	4402	22	AAV09698 PERV env protein c
3	3175.8	95.7	7362	22	AAV09698 PERV env protein c
4	2976.4	89.7	7873	22	AAV09698 PERV env protein c
5	2916.8	87.9	8196	18	AAV09698 Porcine retrovirus
6	2916	87.8	8209	18	AAV09698 Porcine retrovirus
7	2915.2	87.8	7892	18	AAV09698 Porcine retrovirus
8	2915.2	87.8	8132	22	AAV09698 Porcine retrovirus
9	2721.4	82.0	4918	22	AAV09698 Porcine retrovirus
10	2420.6	72.9	8060	18	AAV09698 Porcine retrovirus
11	2420.6	72.9	8060	18	AAV09698 Porcine retrovirus

12	2415.6	72.8	6076	22	AAV09698
13	2414	72.7	6076	22	AAV09698
14	1771.8	53.4	7333	22	AAV09698
15	1771.8	53.4	7333	22	AAV09698
16	1175.8	35.4	8655	20	AAV09698
17	1131	34.1	3482	20	AAV09698
18	1088	32.8	8088	21	AAV09698
19	1088	32.8	8088	21	AAV09698
20	941.4	28.4	8278	22	AAV09698
21	940	28.3	8892	16	AAV09698
22	933.6	28.1	8323	16	AAV09698
23	933.6	28.1	8323	16	AAV09698
24	933.6	28.1	8323	16	AAV09698
25	932	28.1	8323	15	AAV09698
26	932	28.1	10367	15	AAV09698
27	932	28.1	10367	18	AAV09698
28	932	28.1	10367	18	AAV09698
29	916.2	27.6	3643	21	AAV09698
30	916.2	27.6	5178	21	AAV09698
31	916.2	27.6	7308	18	AAV09698
32	916.2	27.6	7308	18	AAV09698
33	916.2	27.6	7616	18	AAV09698
34	916.2	27.6	8330	20	AAV09698
35	916.2	27.6	8332	17	AAV09698
36	916.2	27.6	8332	18	AAV09698
37	916.2	27.6	8332	19	AAV09698
38	916.2	27.6	8332	19	AAV09698
39	916.2	27.6	8332	19	AAV09698
40	916.2	27.6	8332	21	AAV09698
41	916.2	27.6	8332	21	AAV09698
42	916.2	27.6	8332	21	AAV09698
43	916.2	27.6	8332	21	AAV09698
44	916.2	27.6	12135	18	AAV09698
45	916.2	27.6	36538	18	AAV09698

# ALIGNMENTS

RESULT 1	AAV09698	standard; DNA; 3320 BP.
ID	AAV09698	
AC	AAV09698	
XX		
DT	19-MAY-1998	(first entry)
XX		
DE	Porcine retrovirus pol and env DNA.	
XX		
KW	Porcine retrovirus; POEV; POL protein; ENV protein; vaccine;	
KW	diagnosis; xenotransplantation; prophylactic; therapeutic; ds.	
XX		
OS	Porcine retrovirus.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	23..2793
FT		/*tag= a
FT		/product= POL protein
FT		/note= "polymerase protein"
FT	CDS	2642..3297
FT		/*tag= b
FT		/product= ENV protein
FT		/note= "envelope protein"
XX		
XX	W09740167-A1.	
XX		
XX	30-OCT-1997.	
XX		
XX	18-APR-1997;	97W0-GB01087.
XX		
XX	10-FEB-1997;	97GB-0002668.
XX	19-APR-1996;	96GB-0008164.
XX		

PERV env protein c  
PERV env protein c  
Defective retrovir  
Porcine retrovirus  
Mus dunni endogeno  
Pig endogenous ret  
Complete nucleotid  
Galy SEAO genome.  
Canine retrovirus  
Osteoinductive ret  
Friend murine leuk  
LTR clone of PB29  
Friend ecotopic m  
LTR clone of PB29  
PLR8332. Syntheti  
Sequence of plasm  
Glycopeptide expe  
MLV reverse transc  
Gibbon leukemia v  
Plasmid hCMV+Intro  
Plasmid Ceb coding  
MOMV gene sequenc  
MOMV gag/pol gene  
MOMV murine leu  
MOMV murine leu  
Retroviral vector  
Retroviral vector  
Retroviral gag gen  
MMLV gag, gag-pol  
MOMV murine leu  
Complete nucleotid  
Plasmid pADCMVgag-  
Recombinant trans-



QY	1801	gttcctcaggtaccagagtggtctgactcgtgtgtc	aaacatctgtgtgccttcgcagctg	1860
Db	1801	gtctcagagctaccagagtggtctgactcgtgtgtc	aaacatctgtgtgccttcgcagctg	1860
QY	1861	gttaatgctaaatccttcacgaatacctccagaaagag	atcaagggagagccatccacgc	1920
Db	1861	gttaatgctaaatccttcacgaatacctccagaaagag	atcaagggagagccatccacgc	1920
QY	1921	gtccactcggaaatgtgacttcactctaggttaaaagcc	gggttaataacggaaacaaatctct	1980
Db	1921	gtccactcggaaatgtgacttcactctaggttaaaagcc	gggttaataacggaaacaaatctct	1980
QY	1981	tttgtctttgttagacaccttttcaggaatggttagag	gcttctactaaagaaagagact	2040
Db	1981	tttgtctttgttagacaccttttcaggaatggttagag	gcttctactaaagaaagagact	2040
QY	2041	tcaacgctgtgtctgaagaataactctgagagaaat	ttttccaagatttgtgatactccaa	2100
Db	2041	tcaacgctgtgtctgaagaataactctgagagaaat	ttttccaagatttgtgatactccaa	2100
QY	2101	gtatacgggtgcacacaaatgtctccagcttctgtcc	cggtgaagtcacagagctgcacaa	2160
Db	2101	gtatacgggtgcacacaaatgtctccagcttctgtcc	cggtgaagtcacagagctgcacaa	2160
QY	2161	atatctggggatttgatgtgaaactctgatatgtgtg	atacagaccccaagctccagaca	2220
Db	2161	atatctggggatttgatgtgaaactctgatatgtgtg	atacagaccccaagctccagaca	2220
QY	2221	gagaagatgataagaacctttaagaagaccttaccaa	atttgaccacagagagactgcat	2280
Db	2221	gagaagatgataagaacctttaagaagaccttaccaa	atttgaccacagagagactgcat	2280
QY	2281	aagatctgtgacgtctccctccgtcccttgtgcttt	taagtgtagaagaaacccctgagac	2340
Db	2281	aagatctgtgacgtctccctccgtcccttgtgcttt	taagtgtagaagaaacccctgagac	2340
QY	2341	tttggtgcgtgacccccaatgatctgctctacaggg	gggagcccccoccgctgtgcagaaat	2400
Db	2341	tttggtgcgtgacccccaatgatctgctctacaggg	gggagcccccoccgctgtgcagaaat	2400
QY	2401	tttggacataatgtctgatatgtctgcgtcttccac	gcttctgtctctcctcctcctcctc	2460
Db	2401	tttggacataatgtctgatatgtctgcgtcttccac	gcttctgtctctcctcctcctcctc	2460
QY	2461	gagtgggtgagggacagacgtctgaaagccagcc	cccgagagccctactctcagggagagact	2520
Db	2461	gagtgggtgagggacagacgtctgaaagccagcc	cccgagagccctactctcagggagagact	2520
QY	2521	caagttccacatcgcgtctccaagttctgagaat	ctcagtcctatgttagacgcacacgt	2580
Db	2521	caagttccacatcgcgtctccaagttctgagaat	ctcagtcctatgttagacgcacacgt	2580
QY	2581	aaactcggagctcgtgtggaagggagacttatctc	gttaacttttggacacacaaagcgt	2640
Db	2581	aaactcggagctcgtgtggaagggagacttatctc	gttaacttttggacacacaaagcgt	2640
QY	2641	aaagtcgaaagaaatcccccacccctgtgatactca	tgatcccaagttaaagccggcgcaac	2700
Db	2641	aaagtcgaaagaaatcccccacccctgtgatactca	tgatcccaagttaaagccggcgcaac	2700
QY	2701	gattcgggtgtgaaagccgaaagagctgaaatcc	ccctctaaagctcgcctccatcgcgt	2760
Db	2701	gattcgggtgtgaaagccgaaagagctgaaatcc	ccctctaaagctcgcctccatcgcgt	2760
QY	2761	gttccttaacctgtcaataaactcctccaagttaa	tgtaagccgtctgtgcagcccgaa	2820
Db	2761	gttccttaacctgtgtcaataaactcctccaagttaa	tgtaagccgtctgtgcagcccgaa	2820
QY	2821	ctcccaataaacccctatctctcaacccgtgttact	taactgactccgtgtacaagttataat	2880
Db	2821	ctcccaataaacccctatctctcaacccgtgttact	taactgactccgtgtacaagttataat	2880
QY	2881	taacagacactcaagggagagctccctcttgggagac	ctgtgtgccttgatattatgtctgc	2940

Db	2881	taaacgaccccaaggaggccctcccttggggaacctgggacctgaaatataatgtctgcct	2940
Oy	2941	tcgaacagtaatccctggtctcaatgacacagccacaccccccgatctacccgtctta	3000
Db	2941	tcgacacgtaatccctggtctcaatgacacagccacaccccccgatctacccgtctta	3000
Oy	3001	cgggttttaagtttggccacgagaccccccaataatgaaagatatgtggaaatccctcaga	3060
Db	3001	cgggttttaagtttggccacgagaccccccaataatgaaagatatgtggaaatccctcaga	3060
Oy	3061	tttcctttgcaagaatgtagctgtcatacttcaatgatatgggaattggaaatggcccat	3120
Db	3061	tttcctttgcaagaatgtagctgtcatacttcaatgatatgggaattggaaatggcccat	3120
Oy	3121	cttcacgaagaagaagtagtaagtactctttgtttaaacaatccacagttataataat	3180
Db	3121	cttcacgaagaagaagtagtaagtactctttgtttaaacaatccacagttataataat	3180
Oy	3181	taattatgagcatcgggagatgggaaagatctggcaacagcggtacaaaaagatgacgaa	3240
Db	3181	taattatgagcatcgggagatgggaaagatctggcaacagcggtacaaaaagatgacgaa	3240
Oy	3241	taagcaataaagctgtcattcgtttgagacctagattacttaaaaataagtttacttaaaa	3300
Db	3241	taagcaataaagctgtcattcgtttgagacctagattacttaaaaataagtttacttaaaa	3300
Oy	3301	aaaaaaaaaaaaaaaaaaaaa 3320	
Db	3301	aaaaaaaaaaaaaaaaaaaaa 3320	
RESULT 2			
AAC67020			
ID	AAC67020 standard; DNA; 4402 BP.		
XX	AAC67020;		
AC			
XX	27-MAR-2001 (first entry)		
DT			
XX	PERV env protein coding sequence SEQ ID NO: 20.		
DE			
XX	Xenotransplantation; infectious agent; vaccine; ds.		
KW			
XX	Porcine endogenous retrovirus.		
OS			
XX	MO200071726-A1.		
PN			
XX	30-NOV-2000.		
PD			
XX	24-MAY-2000; 2000MO-US14296.		
PF			
XX	24-MAY-1999; 99US-0135631.		
PR			
XX	(MAYO-) MAYO MEDICAL VENTURES.		
PA			
XX	Federspiel MJ;		
PI			
XX	WPI: 2001-032041/04.		
DR			
XX	Inhibiting or preventing infectious agent transmission in mammalian		
PT	transplant recipients, by introducing recombinant DNA comprising DNA		
PT	encoding extracellular proteins of the agent into donor cells, such as		
PT	swine cells -		
XX			
PS	Claim 16; Page 105-106; 144pp; English.		
CC	The present invention provides a method to prevent the transmission of		
CC	infectious agents during xenotransplantation. This involves introducing		
CC	to donor swine cells a recombinant DNA encoding a peptide fragment from		
CC	the infectious agent, and then introducing these cells into the		
CC	transplant recipient.		
XX			

SO Sequence 4402 BP: 1259 A; 1085 C; 1111 G; 947 T; 0 other:

Query Match 95.8%; Score 3179; DB 22; Length 4402;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 3231; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 22 gatgctctctctgcttgagatgaaccccaatgagccaaacttttgccttgaatg 81  
DB 1034 gatgcctttctgcttgagatgaaccccaatgagccaaacttttgccttgaatg 1093  
QY 82 agagatccaggtacgggaagaacccggcagctaacctcgagcccgagcccaagggttc 141  
DB 1094 agagatccaggtacgggaagaacccggcagctaacctcgagcccgagcccaagggttc 1153  
QY 142 aagaactcccgacatctttgacgaagccctacacagaggagccctggcacaattcagatc 201  
DB 1154 aagaactcccgacatctttgacgaagccctacacagaggagccctggcacaattcagatc 1213  
QY 202 caacacctcaggtgacccctccctccagtagtgatgacctgtcttgcggagagccacc 261  
DB 1214 caacacctcaggtgacccctccctccagtagtgatgacctgtcttgcggagagccacc 1273  
QY 262 aaaaagagctgttagaagtagaagtagaagtagaagtagaagtagaagtagaagtagaag 321  
DB 1274 aaaaagagctgttagaagtagaagtagaagtagaagtagaagtagaagtagaagtagaag 1333  
QY 322 aagaacctctgtaagaagcgccagatcttcagagagaggttaacatacttgggggtacat 381  
DB 1334 aagaacctctgtaagaagcgccagatcttcagagagaggttaacatacttgggggtacat 1393  
QY 382 ttgcgggacgggacgagtagctgtagcagagcgacggaagaaactttagtcagatataccg 441  
DB 1394 ttgcgggacgggacgagtagctgtagcagagcgacggaagaaactttagtcagatataccg 1453  
QY 442 gcccaacacacgacgcaaaagtagagagaggtttttgggagcagctggatttgcagactg 501  
DB 1454 gcccaacacacgacgcaaaagtagagagaggtttttgggagcagctggatttgcagactg 1513  
QY 502 tggatcccggggtttgacacattagcagcccaactacacggtacccaagaaagaaaggg 561  
DB 1514 tggatcccggggtttgacacattagcagcccaactacacggtacccaagaaagaaaggg 1573  
QY 562 ggaattcctctggtcctctgagcagcaggaaggaattttagtctatacaaaagggcctctg 621  
DB 1574 ggaattcctctggtcctctgagcagcaggaaggaattttagtctatacaaaagggcctctg 1633  
QY 622 agcgacactgtcttgccctccctctgacgttaactaaaccccttaactatgtgtagatgg 681  
DB 1634 agcgacactgtcttgccctccctctgacgttaactaaaccccttaactatgtgtagatgg 1693  
QY 682 cgtlaaggaagtagcccgagaggttttaacccaacaccttagacacagggagagactgtct 741  
DB 1694 cgtlaaggaagtagcccgagaggttttaacccaacaccttagacacagggagagactgtct 1753  
QY 742 gctactctgtcaagaagacttgactctgttagcagatggtttggccgtatgtctgaaggtc 801  
DB 1754 gctactctgtcaagaagacttgactctgttagcagatggtttggccgtatgtctgaaggtc 1813  
QY 802 atgcgagctgtgtagcactactggtcaagagacgctgacaaattgactttgggagacgaatata 861  
DB 1814 atgcgagctgtgtagcactactggtcaagagacgctgacaaattgactttgggagacgaatata 1873  
QY 862 actgataatagcccccaatgacatgtgtagaagacatgcttcggcagcccccaagccagatgagt 921  
DB 1874 actgataatagcccccaatgacatgtgtagaagacatgcttcggcagcccccaagccagatgagt 1933  
QY 922 accaagcgcgcagtagaaccactatacaagcctgtcttcaacagagaggttcaacttgcgt 981  
DB 1934 accaagcgcgcagtagaaccactatacaagcctgtcttcaacagagaggttcaacttgcgt 1993  
QY 982 ccacgcgcgcgtctcaacctgtgacacttcttgcctggaagagacgtgaaccagtgact 1041  
DB 982 ccacgcgcgcgtctcaacctgtgacacttcttgcctggaagagacgtgaaccagtgact 1041

DB 1994 ccacgcgcgcgtctcaacctgtgacacttcttgcctggaagagacgtgaaccagtgact 2053  
QY 1042 catgattgccaataactattgattgagagacttgggtccggaaggaacttaacacata 1101  
DB 2054 catgattgccaataactattgattgagagacttgggtccggaaggaacttaacacata 2113  
QY 1102 ccgctgacttgaagaagtgtaaccttgcttacttgcaggaagcagatctatgtgtgaaggt 1161  
DB 2114 ccgctgacttgaagaagtgtaaccttgcttacttgcaggaagcagatctatgtgtgaaggt 2173  
QY 1162 aagaagtagtgcctggggcgagctgtgtgagcgggaacccgacagatcttggccagccgtg 1221  
DB 2174 aagaagtagtgcctggggcgagctgtgtgagcgggaacccgacagatcttggccagccgtg 2233  
QY 1222 ccggaaggaacttcaacgcaaaagcttgagctcagtcagccctcagcaagctttgcgctg 1281  
DB 2234 ccggaaggaacttcaacgcaaaagcttgagctcagtcagccctcagcaagctttgcgctg 2293  
QY 1282 gccgaagggaaatccataaacaattlaacgcaacgagtagtgccttgcagctcacac 1341  
DB 2294 gccgaagggaaatccataaacaattlaacgcaacgagtagtgccttgcagctcacac 2353  
QY 1342 gtacacggggccatctataaacaaggggtgtgttacttccagcaggaagggaataaag 1401  
DB 2354 gtacacggggccatctataaacaaggggtgtgttacttccagcaggaagggaataaag 2413  
QY 1402 aacaagaagaattcttaagccttatagaagcctttagccaaagaagcttgacttt 1461  
DB 2414 aacaagaagaattcttaagccttatagaagcctttagccaaagaagcttgacttt 2473  
QY 1462 atacactgtccttgcagatcaagaagaacccaagatccatctatctlaaggggaacagatgct 1521  
DB 2474 atacactgtccttgcagatcaagaagaacccaagatccatctatctlaaggggaacagatgct 2533  
QY 1522 gaccgggttgcgaagcagggcagccaggtcgttlaaccttgcctataatagaacggccc 1581  
DB 2534 gaccgggttgcgaagcagggcagccaggtcgttlaaccttgcctataatagaacggccc 2593  
QY 1582 aagaaccccaagacccagacagatgacaccccttagaagaacttggcagaagataaagaata 1641  
DB 2594 aagaaccccaagacccagacagatgacaccccttagaagaacttggcagaagataaagaata 2633  
QY 1642 gaccaggttcttgcagactccggaaggagcctgtctatacctataataggaagaacatcctg 1701  
DB 2654 gaccaggttcttgcagactccggaaggagcctgtctatacctataataggaagaacatcctg 2713  
QY 1702 ccccaaaaagaaggtttagaatattgccaagaataacgttcttaacccacttgaact 1761  
DB 2714 ccccaaaaagaaggtttagaatattgccaagaataacgttcttaacccacttgaact 2773  
QY 1762 aaaaactgtcagaagttgtagaagaatccactatactgttctgagagctacaggggtg 1821  
DB 2774 aaaaactgtcagaagttgtagaagaatccactatactgttctgagagctacaggggtg 2833  
QY 1822 gctgaactcgttgcataaacaattgttgcctctgcagctgtgttaatgttaactcctcaga 1881  
DB 2834 gctgaactcgttgcataaacaattgttgcctctgcagctgtgttaatgttaactcctcaga 2893  
QY 1882 atacctcagaagaagaagactaaggggaagccacccagcgctcaactcgtgggaagtgacttc 1941  
DB 2894 atacctcagaagaagaagactaaggggaagccacccagcgctcaactcgtgggaagtgacttc 2953  
QY 1942 acttgagtagaagccggctcaaaatagcgaagaacaatatcattggtttttgttagaaccttt 2001  
DB 2954 acttgagtagaagccggctcaaaatagcgaagaacaatatcattggtttttgttagaaccttt 3013  
QY 2002 tcaagtagtggtagaaggtcttaactactaagaagaagaagactcaaacgtgtgttgcgaagaa 2061  
DB 3014 tcaagtagtggtagaaggtcttaactactaagaagaagaagactcaaacgtgtgttgcgaagaa 3073  
QY 2062 atacttgagaagaatttttccaagaatttgcgaatcttaagtagaagtaggtagaacaatgtgt 2121  
DB 3074 atacttgagaagaatttttccaagaatttgcgaatcttaagtagaagtaggtagaacaatgtgt 3133

QY 2122 ccagcttcgttgcgcagtaagtcaaggactggccaaatatgtggtatgttgaataa 2181  
 |||||  
 Db 3134 ccagcttcgttgcgcagtaagtcaaggactggccaaatatgtggtatgttgaataa 3193  
 QY 2182 ctgcatgtgtacacagaccacaaagctcagagacagtagagagatgaataaacat 2241  
 |||||  
 Db 3194 ctgcatgtgtacacagaccacaaagctcagagacagtagagagatgaataaacat 3253  
 QY 2242 aaagaagaccttaccacaaatggacacagacagtgacattatgttggatgctcttc 2301  
 |||||  
 Db 3254 aaagaagaccttaccacaaatggacacagtgacattatgttggatgctcttc 3313  
 QY 2302 ccccttggctttttagagtgagagaaacacccctggacagatttggtgacccctatga 2361  
 |||||  
 Db 3314 ccccttggctttttagagtgagagaaacacccctggacagatttggtgacccctatga 3373  
 QY 2362 ttgctcgaagggagaccccccgttggcagaatattgcttgcacatagtgtcatgt 2421  
 |||||  
 Db 3374 ttactcgaagggagaccccccgttggcagaatattgcttgcacatagtgtcatgt 3433  
 QY 2422 ctgcttcacagaccttggcttctcttaggtcgaagcgctgagtggtgagcgagcg 2481  
 |||||  
 Db 3434 ctgcttcacagaccttggcttctcttaggtcgaagcgctgagtggtgagcgagcg 3493  
 QY 2482 tggagagagctccggagagcctactc--agagagagacttgcaagttccacatgcctc 2538  
 |||||  
 Db 3494 tggagagagctccggagagcctactcagagagagagacttgcaagttccacatgcctc 3553  
 QY 2539 caagttgagagattcagctcttlttagagcgccacgctgcaagaaacctcgagctcg 2598  
 |||||  
 Db 3554 caagttgagagattcagctcttlttagagcgccacgctgcaagaaacctcgagctcg 3613  
 QY 2599 aaggagaccttaccctgacttcttgacacacaaacggctgtgaagttcgaagaaatccc 2658  
 |||||  
 Db 3614 aaggagaccttaccctgacttcttgacacacaaacggctgtgaagttcgaagaaatccc 3673  
 QY 2659 acctggtacatgcacatccacgcttaagcgcgccacacctccgattcgggtgtgaagcc 2718  
 |||||  
 Db 3674 acctggtacatgcacatccacgcttaagcgcgccacacctccgattcgggtgtgaagcc 3733  
 QY 2719 gaaaagctgaataatcccttaagcttgcctccatccgctggtgttcccttaacttgc 2778  
 |||||  
 Db 3734 gaaaagctgaataatcccttaagcttgcctccatccgctggtgttcccttaacttgc 3793  
 QY 2779 aactcccaagtttaagtgaacgacctgtgtgacagcccaactccatcaaaccttacc 2838  
 |||||  
 Db 3794 aactcccaagtttaagtgaacgacctgtgtgacagcccaactccatcaaaccttacc 3853  
 QY 2839 tctcaacctgttacttactactcagctgtagaagttatataatlaaacagcactcaaggg 2898  
 |||||  
 Db 3854 tctcaacctgttacttactactcagctgtagaagttatataatlaaacagcactcaaggg 3913  
 QY 2899 ggtctcccttgggagactggtggtgacataatattgtctgcttgcagtaactccctg 2958  
 |||||  
 Db 3914 ggtctcccttgggagactggtggtgacataatattgtctgcttgcagtaactccctg 3973  
 QY 2959 tctcaatgaccagggccacaccccgatgactcgtgtcttaagggttttaagttggccc 3018  
 |||||  
 Db 3974 tctcaatgaccagggccacaccccgatgactcgtgtcttaagggttttaagttggccc 4033  
 QY 3019 aggaaccccccaataatgaagaatatgtgtgaatatccctcaggaatttcccttgcagaag 3078  
 |||||  
 Db 4034 aggaaccccccaataatgaagaatatgtgtgaatatccctcaggaatttcccttgcagaag 4093  
 QY 3079 gaactgataacttcttaagtgaatgtgaatgtgaatgtgcacgtctcagcaagaagag 3138  
 |||||  
 Db 4094 gaactgataacttcttaagtgaatgtgaatgtgaatgtgcacgtctcagcaagaagag 4153  
 QY 3139 aagttactcttcttgaataacatcctacaggttaataatlaattgaagcagaggg 3198  
 |||||  
 Db 4154 aagttactcttcttgaataacatcctacaggttaataatlaattgaagcagaggg 4213

QY 3199 atggaagaattggcaacagcggttacaataaagatgttacaataatgaagctgtca 3258  
 |||||  
 Db 4214 atggaagaattggcaacagcggttacaataaagatgttacaataatgaagctgtca 4273  
 QY 3259 ttgcttagaccttagattacttaaaaaaagtttccactaaataaaaaa 3317  
 |||||  
 Db 4274 ttgcttagaccttagattacttaaaaaaagtttccactgaaataaagaaacagaataa 4332

## RESULT 3

AAC67019  
 ID AAC67019 standard; DNA; 7362 BP.

AC AAC67019;

XX 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 19.

XX Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

PN WO200071726-A1.

PD 30-NOV-2000.

PF 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ.

DR WPI: 2001-032041/04.

XX

PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

PT swine cells -

PS Claim 16; Page 101-104; 144pp; English.

CC The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

XX

SQ

Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

## Query Match

95.7%; Score 3175.8; DB 22; Length 7362;  
 Best Local Similarity 97.9%; Pred. No. 0;  
 Matches 3229; Conservative 0; Mismatches 67; Indels 3; Gaps 1;

QY 22 gatgctctcttctgctgagattaaacccacagacacacttttggcttgaatgg 81  
 |||||  
 Db 3633 gatgctcttcttgcctgagattaaacccacagacacacttttggcttgaatgg 3692  
 QY 82 aagatccaggtgaggaagaacgggagctcacttgagccgactggcccaagggttc 141  
 |||||  
 Db 3693 aagatccaggtgaggaagaacgggagctcacttgagccgactggcccaagggttc 3752  
 QY 142 aagaactcccgacacatcttgaaggaagccctacacagagacctggtgcaactcagatc 201  
 |||||  
 Db 3753 aagaactcccgacacatcttgaaggaagccctacacagagacctggtgcaactcagatc 3812  
 QY 202 caaacccctcaggtgagccctctccagtagtgatgacgtcttctggcgggagcacc 261  
 |||||  
 Db 3813 caaacccctcaggtgagccctctccagtagtgatgacgtcttctggcgggagcacc 3872

QY	262	aacaaagactgtcttaagaagttgaagaagcactactgtctgaaattgtctgacctagctac	321
Dp	3873	aaacaaagacvtgcttaagaagaacgaagaagcactactgtctgaaattgtctgacctaaagctac	39322
QY	322	agaagcctctgtcaagaagagccacagatttgcagaagagaggtiaacaactctgtgtgtacagt	381
Dp	3933	agaagcctctgtcaagaagagccacagatttgcagaagagaggtiaacaactctgtgtgtacagt	39922
QY	382	ttcgvggvcgvgcagcgaatgtctgcacgvgagcgcagaaagaaactgtatccagataccg	441
Dp	3993	ttcgvggvcgvgcagcgaatgtctgcacgvgagcgcagaaagaaactgtatccagataccg	40522
QY	442	gcccacacccacagccaaacaaagctgaataagagtttttvggagacgcgtggaattttgcagactg	501
Dp	4053	gcccacacccacagccaaacaaagctgaataagagtttttvggagacgcgtggaattttgcagactg	41122
QY	502	tgtatcccgvggttttcgacaccttagcagccccaactctacccgcttaaccaaagaaaaaggg	561
Dp	4113	tgtatcccgvggttttcgacaccttagcagccccaactctacccgcttaaccaaagaaaaaggg	41722
QY	562	ggaattccctcggggtctctctgtgacacacgaagaaggaatttgatgcatataaaagaagccctgtg	621
Dp	4173	ggaattccctcggggtctctctgtgacacacgaagaaggaatttgatgcatataaaagaagccctgtg	42322
QY	622	agcgacacctgtctctgtgccccctctcgtacgtgaactaaaccccttaaccttatgtgtgaatgag	681
Dp	4233	agcgacacctgtctctgtgccccctctcgtacgtgaactaaaccccttaaccttatgtgtgaatgag	42922
QY	682	cgtaagaagatagcccggaaggaatgtttaaacccaaaccccttagagacgacttgagaagacacctgt	741
Dp	4293	cgtaagaagatagcccggaaggaatgtttaaacccaaaccccttagagacgacttgagaagacacctgt	43522
QY	742	gctcaccctgtccaagaagaagcttgaacctctgtacagcaatgtgtgtgccccgtatgtctgaagct	801
Dp	4353	gctcaccctgtccaagaagaagcttgaacctctgtacagcaatgtgtgtgccccgtatgtctgaagct	44122
QY	802	atcgcaagctgtgcccatactactgtgtcaagagccgtctgaacaaattgacttttggagacagatatata	861
Dp	4413	atcgcaagctgtgcccatactactgtgtcaagagccgtctgaacaaattgacttttggagacagatatata	44722
QY	862	actgtaaatagcccccccaactgtgacttgaagaaacatgtgtctgcagacccccacagaccagatgtagtg	921
Dp	4473	actgtaaatagcccccccaactgtgacttgaagaaacatgtgtgtctgcagacccccacagaccagatgtagtg	45322
QY	922	accaaagcccgcatgtgaccacaactacaagaagcgtctctcaagaagagvgtaacttgcgt	981
Dp	4533	accaaagcccgcatgtgaccacaactacaagaagcgtctctcaagaagagvgtaacttgcgtct	45922
QY	982	ccaaacagccgctctcaaacccctgcacactctctcgtccctgaagaagactgtgaacacagttgact	1041
Dp	4593	ccaaacagccgctctcaaacccctgcacactctctcgtccctgaagaagactgtgaacacagttgact	46522
QY	1042	catgatgtcatcaactatgtatgttgaagagagctvggttccgaagagaccttacaagacata	1101
Dp	4653	catgatgtcatcaactatgtatgttgaagagagctvggttccgaagagaccttacaagacata	47122
QY	1102	ccggttgacttgagaaagaagtgctaaacacttggttccactgtccagaaagacagtatgtgttgaaggt	1161
Dp	4713	ccggttgacttgagaaagaagtgctaaacacttggttccactgtccagaaagacagtatgtgttgaaggt	47722
QY	1162	aagagvgatgtgtgtgvcgvgcagctgtgtgtgaagacccgcgaacatlttvggacagacagcctg	1221
Dp	4773	aagagvgatgtgtgtgvcgvgcagctgtgtgtgaagacccgcgaacatlttvggacagacagcctg	48322
QY	1222	ccggaagaagaacttcagcgcgaaaaagcgtgcagctcactatggtccctacgcaagcttttgcgctg	1281
Dp	4833	ccggaagaagaacttcagcgcgaaaaagcgtgcagctcactatggtccctacgcaagcttttgcgctg	48922
QY	1282	gccgaaggaatccatcaaaacattttaaagvcagcaagatagctgtgtgtgcactgtgcaactgtcaaac	1341
Dp	4893	gccgaaggaatccatcaaaacattttaaagvcagcaagatagctgtgtgtgcactgtgtgcaactgtcaaac	49522
QY	1342	gtacaacgvggccaactataaacaagaagvgggttgccttaccctcagcagvgagggaaataaag	1401

Db	4953	gtacacgggccaactcatataagcaaaaggggtctgtcttaacctgaacgaggggaaataaag	5012
Qy	1402	aacaaagaggaatcttaagccttaetagaaccttaacatttgccaanaagsgctagctat	1461
Db	5013	aacaaaggaatcttaagccttaagccttaetagaaccttaacatttgccaanaagsgctagctat	5072
Qy	1462	ataacatgtctcttgatacatcagaagcccaaaagatctcatatctataggggagcaagatgtct	1521
Db	5073	ataacatgtctcttgatacatcagaagcccaaaagatctcatatcgaagggagcaagatgtct	5132
Qy	1522	gaccgggtctgccaagcagcgaccccaagcgctgttaacctctgcctaaetagaacggcc	1581
Db	5133	gaccgggtctgccaagcagcgaccccaagcggtgttaacctctgcctatgtagaaccacc	5192
Qy	1582	aaagcccaagaacccagaacgacagatcacaccttaagaagactgtgcaagagataaaaaata	1641
Db	5193	aaagcccaagaacccagaacgacagatcacaccttaagaagactgtgcaagagataaaaaata	5252
Qy	1642	gaccgggtctcttgagactccgggaagggagacctgtataactctcatataggaagaaatctcg	1701
Db	5253	gaccgggtctcttgagactccgggaagggagacctgtataactctcagatgagggaagaaatctcg	5312
Qy	1702	cccccaaaaagaaggttagaataatgttccaacagatacatcgttctaaccaccaacttagaac	1761
Db	5313	cccccaaaaagaaggttagaataatgttccaacagatacatcgttctaaccaccaacttagaac	5372
Qy	1762	aaacacgtcgagagctgtgtgcagaacatccaccttaacatgtgtcttaagagctaccagaagt	1821
Db	5373	aaacacgtcgagagctgtgtgcagaacatccaccttaacatgtgtcttaagagctaccagaagt	5422
Qy	1822	gctgaactcgtgtgtcacaacatctgtgtcccttcgacgctgtgttaatgtcatctctcaga	1881
Db	5433	gctgaactcgtgtgtcacaacatctgtgtcccttcgacgctgtgttaatgtcatctctcaga	5492
Qy	1882	atacctccaagaagaagaacataaagggggaagccaccggcgctcacactgagggaagtgaatc	1941
Db	5493	atacctccaagaagaagaacataaagggggaagccaccggcgctcacactgagggaagtgaatc	5552
Qy	1942	actgaggttaaaagcccggtctaataacggaacaaatatactatctgttcttgtagacacctt	2001
Db	5553	actgaggttaaaagcccggtctaataacggaacaaatatactatctgttcttgtagacacctt	5612
Qy	2002	tcagagatggtgtagaagggcttatctctactaagaaagaatcttaacogtvgtvgctaaagaa	2061
Db	5613	tcagagatggtgtagaagggcttatctctactaagaaagaatcttaacogtvgtvgctaaagaa	5672
Qy	2062	atactgaggggaatcttctccaagaattgtgatactccaaggttaatagvggtcacaacaatgt	2121
Db	5673	atactgaggggaatcttctccaagaattgtgatactccaaggttaatagvggtcacaacaatgt	5732
Qy	2122	ccaagcttgcttgcaccaagtaagtccaaggaactgccaagaatactvgggagatgttgaa	2181
Db	5733	ccaagcttgcttgcaccaagtaagtccaaggaactgccaagaatactvgggagatgttgaa	5792
Qy	2182	ctgacatgttgataacagaaccccaaggtccaaggtagaagaagatgaataagaacctt	2241
Db	5793	ctgacatgttgataacagaaccccaaggtccaaggtagaagaagatgaataagaacctt	5852
Qy	2242	aaagagaccccttaacaaattgaccaccaagagactcgggataatgatgtgatatgtctctcg	2301
Db	5853	aaagagaccccttaacaaattgaccaccaagagactcgggataatgatgtgatatgtctctcg	5912
Qy	2302	cccttctgtgcttcttlaaggtgaggaagaaaccccttgagcaattggggtcgacccccctatgaa	2361
Db	5913	cccttctgtgcttcttlaaggtgaggaagaaaccccttgagcaattggggtcgacccccctatgaa	5972
Qy	2362	ttgctctacaggggaaccccccgcttgcagaaatgtgctctgcacataagtgctatactg	2421
Db	5973	ttactctacaggggaaccccccgcttgcagaaatgtgctctgcacataagtgctatactg	6032
Qy	2422	ctgctcttccaagccttgtctctcctaggtctaaagcgctcgaatgggtggaagcagcgag	2481



Db	6033	ctgcttcccaagccttgttctctcctcagctctaaagaccttgcgtggttggtgaacaacgacg	6092
QY	2482	tggaaagcagctccggagagcctaactc--aggagagacttgcaggtttccacatcgtctc	2538
Db	6093	tggaaagcaacatcccgagagagctactcagagagagagagacttgcagatccacatcgttcc	6152
QY	2539	caagtttgaagatttcaagctctatgttttaagccacccggtgcagggaaacctgcagatctcgttgcg	2558
Db	6153	caagtttggagatttcaagctctcctcgttttagacgcccacggtgcagagaaacctgcagatctcgttgcg	6212
QY	2559	aagggaccttatctctcgtactcttgcagcacaccacaagcgctgttgaagtcgaagaaatccccc	2658
Db	6213	aagggcccttatccagctacttctgcacacacaaacgagctgtgaagatgcgaagaaatccccc	6272
QY	2659	acctgcgtatccatgtatccccaagcttgaagccggcgccacccctcccgattcgggggtggaagcc	2718
Db	6273	acctgcgtatccatgtatccccaagcttgaagccggcgccacccctcccgattcgggggtggaagcc	6332
QY	2719	gaaagaagctgaaataatccccccttaagcttgcgcctccatccgtgtgtgtctctactctcgtcat	2778
Db	6333	gaaagaagctgaaataatccccccttaagcttgcgcctccatccgtgtgtgtctctactctcgtcat	6332
QY	2779	aacctctcaagtttaatgttaaacgacctgtgtggaacgcccgaacctcccaataaaccttatc	2838
Db	6393	aacctctcaagtttaatgttgaatacgacctgtgtggaacgcccgaacctcccaataaaccttatc	6452
QY	2839	tctcaactctgttacttaacttaactcgcgtacacaggttatataattaaacacatctcaagggga	2888
Db	6453	tctcaactctgttacttaacttgaactcgcgtacacaggttatataattaaacacatctcaagggga	6512
QY	2899	ggctccctctgggagacctgtgtgcctcgtgaattatatagtctgcctctgcatcaagtaaacctctgcg	2958
Db	6513	ggctccctctgggagacctgtgtgcctcgtgaattatatagtctgcctctgcatcaagtaaacctctgcg	6572
QY	2959	tctcaatagaccagagccacaccccccgatgtactccgtgcttaacgggtttacgcttgcctcc	3018
Db	6573	tctcaatagaccagagccacaccccccgatgtactccgtgcttaacgggtttacgcttgcctcc	6632
QY	3019	aggagcccccaaatgaagaatattgtgtggaatctccatgaattctcctttgcaagcaatg	3078
Db	6633	aggagcccccaaatgaagaatattgtgtggaatctccatgaattctcctttgcaagcaatg	6692
QY	3079	gagctgtgataactcttcaatgtatgtggaatttgaatgtgcacgtctctcctcaagcaagaagat	3138
Db	6693	gagctgtgttaactcttcaatgtatgtggaatttgaatgtgcacgtctctcctcaagcaagaagat	6752
QY	3139	aagttactactcttgtttaaacaatccctacacggtataataaatttaattatgtgcattgagag	3198
Db	6753	aagttactactcttgtttaaacaatccctacacggtataataaatttaattatgtgcattgagag	6812
QY	3199	atgtgaaagattgtgcaacacggtgtacaanaaagtgtacgaaataagcaataaagctgtca	3258
Db	6813	atgtgaaagattgtgcaacacggtgtacaanaaagtgtacgaaataagcaataaagctgtca	6872
QY	3259	tctgcttagaccctagattactttaaanaataagtttgccttaanaaaaaaaaaaaaaaaaaaa	3317
Db	6873	tctgcttagaccctagattactttaaanaataagtttgccttgcgtgaaagaaacacagaaata	6931

RESULT	4
AAC67023	
ID	AAC67023 standard; DNA; 7873 BP
XY	

AC	AAC67023;
XX	
DT	27-MAR-2001 (first entry)
XX	

	PERV env protein coding sequence	SEQ ID NO: 23.
DE yy		

Xenotransplantation; infectious agent; vaccine; ds

OS Porcine endogenous retrovirus.

PN	WO2000/1726-A1.
XX	
XX	30-NOV-2000.
PD	
XX	
PF	24-MAY-2000; 2000WO-US14296.
XX	
XX	
PR	24-MAY-1999; 99US-0135631.
XX	
PA	(MAYO-) MAYO MEDICAL VENTURES.
XX	
PI	Federspiel MJ;
XX	
DR	WPI: 2001-032041/04.
XX	
PT	Inhibiting or preventing infectious agent transmission in mammalian
PT	transplant recipients, by introducing recombinant DNA comprising DNA
PT	encoding extracellular proteins of the agent into donor cells, such as
PI	swine cells -
XX	
PS	Claim 16; Page 112-115; 14pp; English.
XX	
CC	The present invention provides a method to prevent the transmission of
CC	infectious agents during xenotransplantation. This involves introduc
CC	to donor swine cells a recombinant DNA encoding a peptide fragment fro
CC	the infectious agent, and then introducing these cells into the
CC	transplant recipient.
XX	
SD	Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

Query Match	89.7%	Score	2976.4	DB 22	Length	7873
Best Local Similarity	95.1%	Pred	No. 0			
Matches 3139	Conservative	0	Mismatches	71	Indels	90
					Gaps	3

OY	22	gaacgtttcttcgtccgagattacaccccaactagccaaacactttttgcttcgtagtg	81
Db	2447	gattgcctttcttcgtccgagattacaccccaactagccaaacactttttgcttcgtagtg	2506
OY	82	agagatccaggtacaggaagaagaacccggacagctacactctggaccggactgtgccaaaggttc	141
Db	2507	agagatccaggtacaggaagaagaacccggacagctacactctggaccggactgtgccaaaggttc	2566
OY	142	aagaactccccgaccatcttttgacgaagccctacacaaagagacctgtgcccacttcagatc	201
Db	2567	aagaactccccgaccatcttttgacgaagccctacacaaagagacctgtgcccacttcagatc	2626
OY	202	caaacacctgaagtgtagaccctccctccagtagatgtgatgcactgtctctgtgcggagacac	261
Db	2627	caaacacctgaagtgtagaccctccctccagtagatgtgatgcactgtctctgtgcggagacac	2686
OY	262	aaacagagactgtcttagaagaagtacgaagacactactcttggaattgtctgaacttagctac	321
Db	2687	aaacagagactgtcttagaagaagaagacgaagacactactctgtgaattgtctgaacttagctac	2748
OY	322	agagactctgtctaaagaagcccaagattctgcagagagaagatacaactctgtgtgtacagt	381
Db	2747	agagactctgtctaaagaagcccaagattctgcagagagaagatacaactctgtgtgtacagt	2806
OY	382	tttgagggtgcggagcgaatgtgtcgaagcgggacacggaagaanaaacgttagtccagatcacg	441
Db	2807	tttgagggtgcggagcgaatgtgtcgaagcgggacacggaagaanaaacgttagtccagatcacg	2866
OY	442	gccccaacccaagccaacaaagctgaagaagatttttgaggacagcgtgattttgcagactg	501
Db	2867	gccccaacccaagccaacaaagctgaagaagatttttttgggacagcgtgattttgcagactg	2926
OY	502	tgatgcccggttttgagaccttagcagccccaactacccgtctaaccaagaanaaagg	561
Db	2927	tgatgcccggttttgagaccttagcagccccaactacccactaaccaagaanaaagg	2986
OY	562	ggaattccctgcggctctctagacccaaggaagcattgtatgtcatataaanaagccctgtg	621
Db	2987	ggaattccctgcggctctctagacccaaggaagcattgtatgtcatataaanaagccctgtg	3046



Db	5121	taactccctaaggttaactatggttaaaagcctctgttgacacgcccgaactcccataaaccttat	5180
QY	2838	cttcaacctggttaacttaacttaactgaactcoggtacaaagtataataataacagcactcaaggg	28937
Db	5181	ctctcaacttggcttaacttaactgaactcoggtacagtgattataataataacagcactcaaggg	5240
QY	2898	aggtccctctggggagcacttggcttgccgtgaatatatgtctgcgcttgcatacgaatccctg	2957
Db	5241	aggccctcttggggagccttggcttgccgtgaatatatgtctgcgcttgcatacgaatccctg	5300
QY	2958	gtctcaatgacaaagcgacaccccccccgatgtactccgtctctaaaggttttaagcttgc	3017
Db	5301	gtctcaatgacacagcgacaccccccccgatgtactccgtctctaaaggttttaagcttgc	5360
QY	3018	caggaccccccaataataatgaagaatatgttggaaatccctcagaattctctttgcaagcaat	3077
Db	5361	caggaccccccaataataatgaagaatatgttggaaatccctcagaattctctttgcaagcaat	5420
QY	3078	ggagctgcataactcttaacttaagatgatgggaatttggaatggcgactcctcaagcaagacag	3137
Db	5421	ggagctgcgttaactcttaacttaagatgatgggaatttggaatggcgactcctcaagcaagacag	5480
QY	3138	taagttaactcttcttgttaacaatccctcaccaggtataatcaattcaattatgtgcgatgga	3197
Db	5481	taagttaactcttcttgttaacaatccctcaccaggtataatcaattcaattatgtgcgatgga	5540
QY	3198	gatgtgaagaatgtggcaacagcgggtacaacaagaatgtlagaataaagaacaaataagctgtc	3257
Db	5541	gatgtgaagaatgtggcaacagcgggtacaacaagaatgtlagaataaagaacaaataagctgtc	5600
QY	3258	atctggttagacccagattactttaaaaaataagtttaacttaaaaaataaaaaataaaaaa	3317
Db	5601	atctggttagacccagattactttaaaaaataagtttaacttaaaaaataaaaaataaaaaa	5660

RESULT	5
AAV09699	ID AAV09699 standard; DNA; 8196 BP.
XX	
XX	AAV09699;
XX	
DT	19-MAY-1998 (first entry)
XX	
DE	Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX	
KW	Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
KM	vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds
XX	
OS	Porcine retrovirus.
XX	
FH	Key Location/Qualifiers
FH	576..2126
FT	CDS
FT	/tag= a
FT	/product= GAG protein
FT	/note= "Viron core polypeptide"
FT	2143..5733
FT	/tag= b
FT	/product= POL protein
FT	/note= "polymerase peptide sequence as given in
FT	specification"
FT	5606..7576
FT	/tag= c
FT	/product= ENV protein
FT	/note= "envelope protein"
XX	
PN	WC9740167-A1.
XX	
PD	30-OCT-1997.
XX	
PF	18-APR-1997; 97WO-GB01087.
XX	
XX	10-FEB-1997; 97GB-0002668.
RR	

PR	19-APR-1996;	96GB-0008164.
XX		
PA	(IMOT-)	IMOTRAN LTD.
PA	(OONE-)	Q-OONE BIOTECH LTD.
XX		
PI	Galbraith DN,	Haworth C, Lees GM, Smith KT;
XX		
DR	WPI; 1997-535851/49.	
XX		
PT	Polynucleotide encoding porcine retrovirus expression product -	
PT	useful to develop products for use in vaccines, diagnosis and	
PT	xeno-transplantation	
XX		
PS	Claim 4; Fig 2; 69pp.	English.
XX		
XX	This DNA sequence encodes the porcine retrovirus (POEV) virion core	
CC	polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These	
CC	proteins can be used to develop viral vaccines, antisense nucleic acids,	
CC	ribozymes and other antiviral agents. They can also be used in	
CC	xeno-transplantation technology and as diagnostic tools.	
XX		
SO	Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;	

Query Match	87.9%	Score 2916.8	DB 18	Length 8196
Best Local Similarity	96.0%	Pred. No. 0		
Matches 3013	Conservative	2	Mismatches 114	Indels 9
				Gaps 2

QY	19	acaaagctgcttcttgcgttgagattaaaccccaactagcaaacacacttttgccttcgaa	78
Db	2956	aaagaaagcctcttcttcgcgttgagattaaaccccaactagcaaacacacttttgccttcgaa	3015
QY	79	tggagagatccacggtgacgggaaagaaacccggcagctctacacttgaaaccgagctgcggcaagg	138
Db	3016	tggagagatccacggtgacgggaaagaaacccggcagctctacacttgaaaccgagctgcggcaagg	3075
QY	139	ttcaagaactcccccgaacatcttttgcgaagccctacacaaaggagacttgcgaacttcag	198
Db	3076	ttcaagaactcccccgaacatcttttgcgaagccctacacaaaggagacttgcgaacttcag	3135
QY	199	atccaaaccccttcagatgtacccctctccacgtacgcgtgatgactgtcttcgcgggaagcc	258
Db	3136	atccaaaccccttcagatgtacccctctccacgtacgcgtgatgactgtcttcgcgggaagcc	3195
QY	259	accaaacagagactgtcttgaaaggtacgaagcactactgtcgtgaattgtctgcactaagc	318
Db	3196	accaaacagagactgtcttgaaaggtacgaagcactactactgtgaattgtctgcactaagc	3255
QY	319	tacaaagcctctgcctaaagaagcccaagatttcgaagagaaggttaacactatcttggtgac	378
Db	3256	tacaaagcctctgcctaaagaagcccaagatttcgaagagaaggttaacactatcttggtgac	3315
QY	379	agttgcggggcgcgagcagatgtgcctgtgacccgagacagcagaagaaactgtacgtccagata	438
Db	3316	agttgcggggcgcgagcagatgtgcctgtgacccgagacagcagaagaaactgtacgtccagata	3375
QY	439	ccggcccccaacacacagccaaacaagttggagagatltttggggacagcgtgattttgcaga	498
Db	3376	ccggcccccaacacacagccaaacaagttggagagatltttggggacagcgtgattttgcaga	3435
QY	499	ctgtgatacccccggggtttgcgacactttagcacagcccaactctcccgctaaccaagaagaaaa	558
Db	3436	ctgtgatacccccggggtttgcgacactttagcacagcccaactctacccgcttaaccaagaagaaaa	3495
QY	559	gggggattctctcctgggctctcttgagcacacagaagcatttgatgtatacaaaaaggccctg	618
Db	3496	gggggattctctcctgggctctcttgagcacacagaagcatttgatgtatacaaaaaggccctg	3555
QY	619	ctgagcgacactgtcctctgcgcctcccttaagtaactaaaccccttaaccccttatgtgat	678
Db	3556	ctgagcgacactgtcctctgcgcctcccttaagtaactaaaccccttaaccccttatgtgat	3615
QY	679	gagcgtlaaggagtagcccgagagatttlaacccaacacacttagacacatlgagagagacct	738

Db 3616 gaagctaaaggagtagccagagagatttaaccnaaaccttagaccatgagagacct 3675  
Qy 739 gtggcactactgtcaaaagagcttgatcctgtagcagatggttgccctatctcgaag 798  
Db 3676 gtgcctcaactgtcaaaagagcttgacccgttagccagtggttgcccgatactccgaag 3735  
Qy 799 gctatcgcaactgtgscatctactgtcaaaagagcttgacaaatctgactttggagcagaat 858  
Db 3736 gctatcgcaactgtgscatctactgtcaaaagagcttgacaaatctgactttggagcagaat 3795  
Qy 859 ataactgtaatacccccccttgatcttgagaaacatcgcttcgagcccccagacgaatg 918  
Db 3796 ataactgtaatacccccccttgatcttgagaaacatcgcttcgagcccccagacgaatg 3855  
Qy 919 atgacaacgcccgcatacgaacacacttcaaacgctgtcttcacagagaggtgctacttc 978  
Db 3856 atgacaacgcccgcatacgaacacacttcaaacgctgtcttcacagagaggtgctacttc 3915  
Qy 979 gtccacacagcgcctcacaacccctgcacactctctgcctgaagagactgaatgaacagtg 1038  
Db 3916 gtccacacagcgcctcacaacccctgcacactctctgcctgaagagactgaatgaacagtg 3975  
Qy 1039 actcatgatgcccatacactatgattgtagagagactggtggtccgcaagacacttaacagc 1098  
Db 3976 actcatgatgcccatacactatgattgtagagagactggtggtccgcaagacacttaacagc 4035  
Qy 1099 ataccgctgactgtgagaagtgcttaacctgttctactgacgaaagcagctatggtgtgaa 1158  
Db 4036 ataccgctgactgtgagaagtgcttaacctgttctactgacgaaagcagctatggtgtgaa 4095  
Qy 1159 ggttaagagagatggtcgtggcgcgagctggtgagacgggacccgacagacttgagccagcagc 1218  
Db 4096 ggttaagagagatggtcgtggcgcgagctggtgagacgggacccgacagacttgagccagcagc 4155  
Qy 1219 ctgcacgaaaggaacttcaagcgcaaaagcgctgaactatgagccctcaacgaaagcttgccg 1278  
Db 4156 ctgcacgaaaggaacttcaagcgcaaaagcgctgaactatgagccctcaacgaaagcttgccg 4215  
Qy 1279 ctgacgaaaggaacttcaacataacatttatacgaacagcaggtatgacttctgagactga 1338  
Db 4216 ctgacgaaaggaacttcaacataacatttatacgaacagcaggtatgacttctgagactga 4275  
Qy 1339 caagctaacggtgcatctatacaaaaggggttgcttaactcagcagagggaggaata 1398  
Db 4276 caagctaacggtgcatctatacaaaaggggttgcttaactcagcagagggaggaata 4335  
Qy 1399 aagacaagaaggaatcttaagcctattagaaagccttaccatttgcacaaaggtactgct 1458  
Db 4336 aagacaagaaggaatcttaagcctattagaaagccttaccatttgcacaaaggtactgct 4395  
Qy 1459 attatacactgtccctggaacatcagaagaccgaatcctacatctagagaggaaccagtg 1518  
Db 4396 attatacactgtccctggaacatcagaagaccgaatcctacatctagagaggaaccagtg 4455  
Qy 1519 gctgacccggtgttgcacagcagcgagcccgagcgctgtaaaccttctgcttaataagaag 1578  
Db 4456 gctgacccggtgttgcacagcagcgagcccgagcgctgtaaaccttctgcttaataagaag 4515  
Qy 1579 ccccaagccccagaaacccagacagctacaccccttagaagactgcaagaagataaanaag 1638  
Db 4516 ccccaagccccagaaacccagacagctacaccccttagaagactgcaagaagataaanaag 4575  
Qy 1639 atagacagtgctctctgagactccggaaggagacgtctatacctcatactggaaggaatc 1698  
Db 4576 atagacagtgctctctgagactccggaaggagacgtctatacctcatactggaaggaatc 4635  
Qy 1699 ctgcccccaaaaagaggttagaataatgctcaacagataacatcgcttaacccacactaga 1758  
Db 4636 ctgcccccaaaaagaggttagaataatgctcaacagataacatcgcttaacccacactaga 4695  
Qy 1759 actaaacacctgagcagtgctgcaaacatccctcatcatgttcttgagctaccaga 1818  
Db 4696 actaaacacctgagcagtgctgcaaacatccctcatcatgttcttgagctaccaga 4755  
Qy 1819 gctgctgactcgtgtgtcaaacatgtgtgtccctgcagcgtgtgttaagtcaatccctcc 1878  
Db 4756 gctgctgactcgtgtgtcaaacatgtgtgtccctgcagcgtgtgttaagtcaatccctcc 4815  
Qy 1879 agaatacctccaagaaagagactaaagggaagccacccagcgctcactctggaagtgaac 1938  
Db 4816 agaatacctccaagaaagagactaaagggaagccacccagcgctcactctggaagtgaac 4875  
Qy 1939 ttccactgtgtaaaagcccggttaaaatacgaagaaataatctatgtgttttgtagaac 1998  
Db 4876 ttccactgtgtaaaagcccggttaaaatacgaagaaataatctatgtgttttgtagaac 4955  
Qy 1999 ttccaagatgtgttagagagctatctactcaataaagaagactcaacccgtgtgtgcaag 2058  
Db 4936 ttccaagatgtgttagagagctatctactcaataaagaagactcaacccgtgtgtgcaag 4995  
Qy 2059 aaaatactggaagaaattttccaaagatttgaatacctaaagtaataggtcagaacat 2118  
Db 4996 aaaatactggaagaaattttccaaagatttgaatacctaaagtaataggtcagaacat 5055  
Qy 2119 ggtccagcttctgttgcacaggttaagtcagagactggtccaaagatatttgagatgttg 2178  
Db 5056 ggtccagcttctgttgcacaggttaagtcagagactggtccaaagatatttgagatgttg 5115  
Qy 2179 aaactgcatctgtcaatacagaccccaagctcagaagctgtagaagagatgaatagaacc 2238  
Db 5116 aaactgcatctgtcaatacagaccccaagctcagaagctgtagaagagatgaatagaacc 5175  
Qy 2239 attaaagagaccccttaacaaatttgaccacagagactgcatatgaattgtagtgcctc 2298  
Db 5176 attaaagagaccccttaacaaatttgaccacagagactgcatatgaattgtagtgcctc 5255  
Qy 2299 ctgccccttgtagcttttagaggtgagaaaccccttgacagcttttgagctaacccctat 2358  
Db 5236 ctgccccttgtagcttttagaggtgagaaaccccttgacagcttttgagctaacccctat 5295  
Qy 2359 gaattgctctaaagggagaccccccctgttgacagaatctgcttgcacatagtgctgat 2418  
Db 5296 gaattgctctaaagggagaccccccctgttgacagaatctgcttgcacatagtgctgat 5355  
Qy 2419 gtgctgcttcccaagcttctctcttaggtctaaaggtcgcaggtgtgtagagcagca 2478  
Db 5356 gtgctgcttcccaagcttctctcttaggtctaaaggtcgcaggtgtgtagagcagca 5415  
Qy 2479 gctgtgaagcagctcccggaagcctactcaagaagagactgcaagttccacatcgctc 2538  
Db 5416 gctgtgaagcagctcccggaagcctactcaagaagagactgcaagttccacatcgctc 5475  
Qy 2539 caagttgagatcagctctatgttagagccagctgacagaaacccctcgagactcggtg 2598  
Db 5476 caagttgagatcagctctatgttagagccagctgacagaaacccctcgagactcggtg 2658  
Qy 2599 aagggaccttactctgtaacttttgacacacacacagcgctgtgaaagtcogaaggaatccc 2658  
Db 5536 aagggaccttactctgtaacttttgacacacacacagcgctgtgaaagtcogaaggaatccc 5555  
Qy 2659 acctggaatcagcttcccaagcttaagcggtgacacccctccggaatcggtgtggaagcc 2718  
Db 5596 acctggaatcagcttcccaagcttaagcggtgacacccctccggaatcggtgtggaagcc 5555  
Qy 2719 gaaaagactgaaatacccttaagcttcgctccatcagctgctgttcttactctgcaat 2778  
Db 5656 gaaaagactgaaatacccttaagcttcgctccatcagctgctgttcttactcttaacat 5715  
Qy 2779 aactcctaagttatgtgtaaacgctgtgagaccccggaactccctaataccctatc 2838  
Db 5716 aactcctaagttatgtgtaaacgctgtgagaccccggaactccctaataccctatc 5775  
Qy 2839 tctcacctgcttacttgaactcgcgtgacagctatataatlaaacagactcaagaa 2898  
Db 5776 tctcacctgcttacttgaactcgcgtgacagctatataatlaaacagactcaagaa 5835

```

QY 2899 ggcctcccttggagacgtgtgacctgaatatatctgctccatcgaatccctcg 2958
      |||||  || ||||| ||||| || ||||| ||||| || |||||
Db 5836 tgcctcctagagcactggtgctcgaactgcatcttcgcctccgatgataaccgcc 5895
      |||||  || ||||| ||||| || ||||| ||||| || |||||
QY 2959 tctcaatgacccagccacaccccccgatgtactccgtcttaagggtttacgttccc 3018
      |||||  || ||||| ||||| || ||||| ||||| || |||||
Db 5896 tglta-----aragcacaccccccacacctagtcgtagtatagggtctatgtctgcc 5949
      |||||  || ||||| ||||| || ||||| ||||| || |||||
QY 3019 aggaccccccaataatgaatattgtggaattctcctcagattccttgcagaatg 3078
      |||||  || ||||| ||||| || ||||| ||||| || |||||
Db 5950 agg-----cacagagaagaagaatactctgagggtctcgggaatctctctgtaagaatg 6006
      |||||  || ||||| ||||| || ||||| ||||| || |||||
QY 3079 gagctgcataacttctaatgatlgtggaattggaatgccaatgctctcagacaagaagat 3138
      |||||  || ||||| ||||| || ||||| ||||| || |||||
Db 6007 gagctgcgtcacctcctcaacagatgagactgtgaatgagccgactctctccagacagcgt 6066
      |||||  || ||||| ||||| || ||||| ||||| || |||||
QY 3139 aagttacctcttctgttaa 3156
      |||||  || ||||| ||||| || ||||| ||||| || |||||
Db 6067 aaattctccttctgcaaa 6084
      |||||  || ||||| ||||| || ||||| ||||| || |||||

RESULT 6
AAV09700
ID AAV09700 standard; DNA; 8209 BP.
AC AAV09700;
XX
XX 20-MAY-1998 (first entry)
DE Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.
XX
KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
KW viron core polypeptide; polymerase protein; envelope protein;
KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
FH
FH Key Location/Qualifiers
FT 1..61 /tag= a
FT misc_feature /note= "R LTR domain"
FT 1..588 /tag= b
FT misc_feature /tag= c
FT 62..143 /tag= c
FT /note= "U5 LTR domain"
FT 588..2162 /tag= d
FT CDS /product= GAG protein
FT /note= "viral core polypeptide"
FT 2163..5747 /tag= e
FT /tag= e
FT /product= POL protein
FT /note= "polymerase peptide"
FT 5620..7590 /tag= f
FT CDS /tag= f
FT /product= ENV protein
FT /note= "envelope protein"
FT 7591..8209 /tag= g
FT LTR /tag= g
FT 7638..8106 /tag= h
FT misc_feature /tag= h
FT /note= "U3 LTR domain"
FT 7643..7686 /tag= i
FT protein_bind /tag= i
FT /note= "Binding site for transcription factors
FT OCT-1, c-Myb, Lys-1, E47 and ETS-1"
FT 7719..7739 /tag= j
FT protein_bind /tag= j
FT /note= "Binding site for transcription factors
FT AP-4 and NF1"
FT 7744..7754 /tag= k
FT /note= "Binding site for transcription factor AP-1/TR"
FT 7773..7785 /tag= l
FT /note= "Binding site for transcription factor ETS-1/GATA"
FT 7793..7834 /tag= m
FT protein_bind /tag= m
FT /note= "Binding site for transcription factors
FT GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT 7841..7851 /tag= n
FT protein_bind /tag= n
FT /note= "Binding site for transcription factor AP-1"
FT 7873..7883 /tag= o
FT protein_bind /tag= o
FT /note= "Binding site for transcription factor AP-1"
FT 7885..7914 /tag= p
FT protein_bind /tag= p
FT /note= "Binding site for transcription factors
FT AP-1/GATA, c-Myb and AP-1"
FT 7928..7953 /tag= q
FT protein_bind /tag= q
FT /note= "Binding site for transcription factors E47
FT and ETS-1"
FT 7964..7968 /tag= r
FT CAAT_signal /tag= r
FT 7985..7997 /tag= s
FT protein_bind /tag= s
FT /note= "Binding site for transcription factor GATA"
FT 8003..8007 /tag= t
FT CAAT_signal /tag= t
FT 8024..8036 /tag= u
FT protein_bind /tag= u
FT /note= "Binding site for transcription factor GATA"
FT 8042..8046 /tag= v
FT CAAT_signal /tag= v
FT 8049..8060 /tag= w
FT protein_bind /tag= w
FT /note= "Binding site for transcription factor
FT AP-1/CREB"
FT 8075..8090 /tag= x
FT TATA_signal /tag= x
FT 1..61 /tag= y
FT misc_feature /tag= y
FT /note= "R LTR domain"
FT 8127..8135 /tag= z
FT misc_feature /tag= z
FT /note= "PolyA downstream element"
FT 8161..8166 /tag= aa
FT /tag= aa

WO9740167-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-GB01087.
XX
XX 10-FEB-1997; 97GB-0002668.
XX 19-APR-1996; 96GB-0008164.
XX
XX (IMOT-) IMOTRAN LTD.
XX (QONE-) Q-ONE BIOTECH LTD.
XX
XX Galbraith DN, Haworth C, Lees GM, Smith KT;
XX WPI: 1997-535851/49.
XX P-PSDB: AAW39271; AAW39272; AAW39273.
XX Polynucleotide encoding porcine retrovirus expression product -
XX useful to develop products for use in vaccines, diagnosis and
XX xeno-transplantation
XX Claim 4; Fig 3; 69pp; English.
XX

```

CC This DNA sequence encodes the porcine retrovirus (POEV) virion core  
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and  
CC also includes the Long Terminal Repeat (LTR). These proteins can be used  
CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
CC antiviral agents. They can also be used in xeno-transplantation  
CC technology and as diagnostic tools.

XX Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other:

Query Match 87.8%; Score 2916; DB 18; Length 8209;

Best Local Similarity 96.0%; Pred. No. 0;

Matches 3014; Conservative 0; Mismatches 115; Indels 9; Gaps 2;

QY 19 aacgaatgctctctctgctggaattacacccactagccaacacttttgccttcga 78  
DB 2970 aagatgctctctctctgctggaattacacccactagccaacacttttgccttcga 3029  
QY 79 tggagatccaggtacggaagaacccggcagctcacctggaaccgactgcccagaag 138  
DB 3030 tggagatccaggtacggaagaacccggcagctcacctggaaccgactgcccagaag 3089  
QY 139 ttcaagaactcccgacacatcttgaacgaacccctacacaggaactggaactcaga 198  
DB 3090 ttcaagaactcccgacacatcttgaacgaacccctacacaggaactggaactcaga 3149  
QY 199 atcaaacctcagtggaacctctccagctacgtggaactgcttctggcggaagcc 258  
DB 3150 atcaaacctcagtggaacctctccagctacgtggaactgcttctggcggaagcc 3209  
QY 259 accaacaagagctgcttaagaaglacgaagcactactgctgaattgtctgacctaagc 318  
DB 3210 accaacaagagctgcttaagaaglacgaagcactactgctgaattgtctgacctaagc 3269  
QY 319 tacaagactcctcttaagaagcccaagattgcaagagaaggtataacttgggtac 378  
DB 3270 tacaagactcctcttaagaagcccaagattgcaagagaaggtataacttgggtac 3329  
QY 379 agtttgcgggcgagcagcagatgctgaacgaagcagcaggaagaactgtagccagata 438  
DB 3330 agtttgcgggcgagcagcagatgctgaacgaagcagcaggaagaactgtagccagata 3389  
QY 439 ccggccccaacacagccaacaaagtgagaagtttttgggacagctggaatttgcaga 498  
DB 3390 ccggccccaacacagccaacaaagtgagaagtttttgggacagctggaatttgcaga 3449  
QY 499 ctgtggaatcccggtgttgcgaaccttagcagcccaactctacccgcttaaccaagaanaa 558  
DB 3450 ctgtggaatcccggtgttgcgaaccttagcagcccaactctacccgcttaaccaagaanaa 3509  
QY 559 ggggattctcctgagctcctgagcaccagaagcattgatctcaaaagggccctg 618  
DB 3510 ggggattctcctgagctcctgagcaccagaagcattgatctcaaaagggccctg 3569  
QY 619 ctgagcgacactgctctgcccctccctgacgttaactaaacctttacacctttatgtgat 678  
DB 3570 ctgagcgacactgctctgcccctccctgacgttaactaaacctttacacctttatgtgat 3629  
QY 679 gagcgttaaggagtagcccgaggaaggtttaaccnaacccctaggaacataaggaagacct 738  
DB 3630 gagcgttaaggagtagcccgaggaaggtttaaccnaacccctaggaacataaggaagacct 3689  
QY 739 gtctcctaactgtcaaaagaagctgtgacccctgttagccagtggttggccgtaattcttaag 798  
DB 3690 gtctcctaactgtcaaaagaagctgtgacccctgttagccagtggttggccgtaattcttaag 3749  
QY 799 gctctcgagctgtgccaactgtgtaagaagcgtgaacaattgtactttggagcagat 858  
DB 3750 gctctcgagctgtgccaactgtgtaagaagcgtgaacaattgtactttggagcagat 3809  
QY 859 ataacttaataagcccccatgcatgtgagaacatcgttcggcagccccagacagctg 918  
DB 3810 ataacttaataagcccccatgcatgtgagaacatcgttcggcagccccagacagctg 3869

QY 919 atgaccagcccgcatgaccacatatacaagcctgtcttcacagagaggtcacttc 978  
DB 3870 atgaccagcccgcatgaccacatatacaagcctgtcttcacagagaggtcacttc 3929  
QY 979 gctccaccagccgctctcaaccctgcaactctctctgctgaagaagctgaatgaaccagt 1038  
DB 3930 gctccaccagccgctctcaaccctgcaactctctctgctgaagaagctgaatgaaccagt 3989  
QY 1039 actatgattgcatcaactatgtatgagaagctgggtgcgcaaggaacttaccagac 1098  
DB 3990 actatgattgcatcaactatgtatgagaagctgggtgcgcaaggaacttaccagac 4049  
QY 1099 ataccgctgactggaagaagtctcaaccctgtctacactgacggaagcagctatgttgaa 1158  
DB 4050 ataccgctgactggaagaagtctcaaccctgtctacactgacggaagcagctatgttgaa 4109  
QY 1159 ggtaaagagatgcttggtgcggaagtggtgaacgggacccgacagatctggccaagc 1218  
DB 4110 ggtaaagagatgcttggtgcggaagtggtgaacgggacccgacagatctggccaagc 4169  
QY 1219 ctgcccgaagaagctctcaagcgaagaagctgaagctcagtcagccacgaagcttgcg 1278  
DB 4170 ctgcccgaagaagctctcaagcgaagaagctgaagctcagtcagccacgaagcttgcg 4229  
QY 1279 ctgcccgaagaagaatccataaacaattatagcgaagcaggtatgcttgcagctgca 1338  
DB 4230 ctgcccgaagaagaatccataaacaattatagcgaagcaggtatgcttgcagctgca 4289  
QY 1339 caagtcacaggggccaactataaagaaggggtgtgcttactactaagcaggaaggaata 1398  
DB 4290 caagtcacaggggccaactataaagaaggggtgtgcttactactaagcaggaaggaata 4349  
QY 1399 aagaacaagaaggaattctaaagcctatataagccttactatgccaagaagctagct 1458  
DB 4350 aagaacaagaaggaattctaaagcctatataagccttactatgccaagaagctagct 4409  
QY 1459 attatacactgctcctgacatcagaagaacccaagaatcctacatctagaagggaacagatg 1518  
DB 4410 attatacactgctcctgacatcagaagaacccaagaatcctacatctagaagggaacagatg 4469  
QY 1519 gctgacccgggttgcgaagcagccagcagctgtttaaacttctgacctataatagaag 1578  
DB 4470 gctgacccgggttgcgaagcagccagcagctgtttaaacttctgacctataatagaag 4529  
QY 1579 cccaagccccgaacccaagacagacagctacacccctagaagactggccaagagataaaaaag 1638  
DB 4530 cccaagccccgaacccaagacagacagctacacccctagaagactggccaagagataaaaaag 4589  
QY 1639 atagacagttctctgagactcctcggaaggagcctgtactactcatatgggaaggaatc 1698  
DB 4590 atagacagttctctgagactcctcggaaggagcctgtactactcatatgggaaggaatc 4649  
QY 1699 ctgcccnaaagaaggggttagaataatgtccaaagatatactgctcaacccaactaga 1758  
DB 4650 ctgcccnaaagaaggggttagaataatgtccaaagatatactgctcaacccaactaga 4709  
QY 1759 actaaacactgtgcaagcagttgtgtcagaacatcccttcatatgttctgaagctaccaga 1818  
DB 4710 actaaacactgtgcaagcagttgtgtcagaacatcccttcatatgttctgaagctaccaga 4769  
QY 1819 gtggtgactggtgtgtaacaatgtgtgtccctgccaagctgtgttaatgtaactcttc 1878  
DB 4770 gtggtgactggtgtgtaacaatgtgtgtccctgccaagctgtgttaatgtaactcttc 4829  
QY 1879 agaatacctccaggaagaagcactaaagggaagcccccagcgcgctcactgtggaatggac 1938  
DB 4830 agaatacctccaggaagaagcactaaagggaagcccccagcgcgctcactgtggaatggac 4889  
QY 1939 ttcaactgagtaagccggtctaaataacgaagaacaaatatactatgtgtttttagaacac 1998  
DB 4890 ttcaactgagtaagccggtctaaataacgaagaacaaatatactatgtgtttttagaacac 4949

QY	1999	tttccagatctggtgagagcttacccttaccataagaagaagcttcaacgcttggtgctaa	2055
Db	4950	ttttccagatctggtgagagcttacccttaccataagaagaagcttcaacgcttggtgctaa	5005
QY	2059	aaaatactctgaggaatttttccagaattctggaatactcaagtgaaatagctcaacaat	2118
Db	5010	aaaatactctgaggaatttttccagaattctggaatactcaagtgaaatagctcaacaat	5065
QY	2119	gttccagcttccgtttgcccagtgtaagctcaagagcttgcacaagatatgtggagttatgg	2178
Db	5070	gttccagcttccgtttgcccagtgtaagctcaagagcttgcacaagatatgtggagttatgg	5129
QY	2179	aaactgcattctgtacatacagaagcccaagaagctcagagatgtagagagtgatatagaac	2238
Db	5130	aaactgcattctgtacatacagaagcccaagaagctcagagatgtagagagtgatatagaac	5188
QY	2239	attaaagagacccttaccacaattgacacagagactgcagattaatgtatgtgactctc	2298
Db	5190	attaaagagacccttaccacaattgacacagagactgcagattaatgtatgtgactctc	5249
QY	2299	ctgcaccttctgccttttaagtgtagaagaacccctgcagagtttgggtctgaccccat	2358
Db	5250	ctgcaccttctgccttttaagtgtagaagaacccctgcagagtttgggtctgaccccat	5309
QY	2359	gaattgctctacggggagaccccccgttgcagaaattgtcccttgcacataatgtctgat	2418
Db	5310	gaattgctctacggggagaccccccgttgcagaaattgtcccttgcacataatgtctgat	5369
QY	2419	gtgcgaccttcccaagcctttgtctcttaagctctaaagcgtctgaatggttgtagagagca	2478
Db	5370	gtgcgaccttcccaagcctttgtctcttaagctctaaagcgtctgaatggttgtagagagca	5429
QY	2479	gctgtgaaagcagctccggagagcctactacccagagagagacttgcgaattccacatgcctc	2538
Db	5430	gctgtgaaagcagctccggagagcctactacccagagagagacttgcgaattccacatgcctc	5489
QY	2539	caagtctggaagctttaaagctctgtttagacgcacacgcgtgcagaggaacctcgagactcgtg	2588
Db	5490	caagtctggaagctttaaagctctgtttagacgcacacgcgtgcagaggaacctcgagactcgtg	5549
QY	2599	aagagaccttactccgcgaactttgacacacccaaacgcgtgtgaaagctcgaagaatcccc	2658
Db	5550	aagagaccttactccgcgaactttgacacacccaaacgcgtgtgaaagctcgaagaatcccc	5609
QY	2659	acctggtacatgcatacccaagctttaaagccggcgacacctcccgattcgggtgtgaaagcg	2718
Db	5610	acctggtacatgcatacccaagctttaaagccggcgacacctcccgattcgggtgtgaaagcg	5669
QY	2719	gaaagagctaaatacccccctaagcttgcgcctccatcgggtgtgtctctactctgtaact	2778
Db	5670	gaaagagctaaatacccccctaagcttgcgcctccatcgggtgtgtctctactctgtaact	5729
QY	2779	aactcctcaagtttaatgtgtaaacgctctgtgacacccgaactcccaataacotlactc	2838
Db	5730	aactcctcaagtttaatgtgtaaacgctctgtgacacccgaactcccaataacotlactc	5789
QY	2839	ttctaacctgttactactgtaagctccggtacaggtatataatatacagacactcaagggga	2888
Db	5790	ttctaacctgttactactgtaagctccggtacaggtatataatatacagacactcaagggga	5849
QY	2899	ggtctcccttggagacctgcgtgcgttaataatataatgtcgtcccttcgatacgaatcccggt	2958
Db	5850	ggtctcccttggagacctgcgtgcgttaataatataatgtcgtcccttcgatacgaatcccggt	5909
QY	2959	ttctaatagcagagccacaccccccgagtgtaactccgtgactcaaggggtttaagtttgcgc	3018
Db	5910	ttctaatagcagagccacaccccccgagtgtaactccgtgactcaaggggtttaagtttgcgc	5963
QY	3019	agagcccccaataatgaagaatgaatgtgtggaactccgaagattccttgcgaaggaat	3078
Db	5964	agagcccccaataatgaagaatgaatgtgtggaactccgaagattccttgcgaaggaat	6020
QY	3079	gagctgcataactcttaatgatgtgggaattggaatactgcagctctctcagcaagaagagt	3138

[illegible]

## RESULT

AAT74884  
ID AAT74884 standard; cDNA; 7892 BP

AC AAT74884

DT 09-FEB-1998 (first entry)  
XY

DE Miniature swine retrovirus cDNA  
XX

kw xenotransplantation: infectious: organ transplantation

[illegible]

XX

XX

FT	585.2159
CDS	

```

ET mat_peptide 585..2156

```

E-I	/note= "pu
ET	0307 57A
CDS	

```
mat pentide 2307 5741
```

FT

$$/ * t a q = e$$

```

ET
-----
/*tag= f

```

XX

XX

XX

XX  
G

XX  
PA (CETIC) A GEN WODTET, OORPXX  
XX  
PT  
Fischman TA.

XX  
DB WP.T. 1997-332804/30

DR F-PSDB; AAW52096-W52096.  
XX

PT viruses in transplant or o

XX

XX

CC containing the putative co

CC (see AAT74812-T74882) can

increase the likelihood of

used to generate probes to

mutations genetic lesions

This cDNA sequence represents a porcine retrovirus from miniature swine containing the putative coding regions for viral GAG, POL and ENV proteins. This sequence and PCR fragments generated from it (see AAT74812-774882) can be used to screen organs for the presence of porcine retroviruses prior to xenotransplantation. Transplantation can increase the likelihood of retroviral activation if intact and infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantation donors. It can be used to detect mutations, genetic lesions or viral recombinants and also to determine

the histological localisation of activated retrovirus. Using Polymerase Chain Reaction DNA Quantitation (PDQ) on blood mononuclear cells, infectivity titration and susceptibility testing can be performed. Ultimately animal donors without intact porcine retroviral sequences or a lower copy number of viral elements could be selected.

Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match 87.8%; Score 2915.2; DB 18; Length 7892;  
Best Local Similarity 94.3%; Pred. No. 0;  
Matches 3051; Conservative 0; Mismatches 173; Indels 12; Gaps 2;

QY 22 gatgacctctctcgtcgtgattatcaacccaccagcagaacacatttttgcctcgaatg 81  
DB 2970 gatgacctctctcgtcgtgattatcaacccaccagcagaacacatttttgcctcgaatg 3029  
QY 82 agagatccaggtacggaagaaacgggagctacactggaacccgactgcccgaagggttc 141  
DB 3030 agagatccaggtacggaagaaacgggagctacactggaacccgactgcccgaagggttc 3089  
QY 142 aagaacctcccgacatctttgacgaagacccctacacagggacttggccaattcagatc 201  
DB 3090 aagaacctcccgacatctttgacgaagacccctacacagggacttggccaattcagatc 3149  
QY 202 caaacacctcaggtgacccctccctccagtaagtgatgacactgcttctgagggagacacc 261  
DB 3150 caaacacctcaggtgacccctccctccagtaagtgatgacactgcttctgagggagacacc 3209  
QY 262 aaacagagactgtctagaaggtacgaaggaactacactgctgaattgtctgacctagctac 321  
DB 3210 aaacagagactgtctagaaggtacgaaggaactacactgctgaattgtctgacctagctac 3269  
QY 322 agagacctctctcgtcgtgattatcaacccaccagcagaacatttgcctcgaatg 381  
DB 3270 agagacctctctcgtcgtgattatcaacccaccagcagaacatttgcctcgaatg 3329  
QY 382 ttgaggggagcgagcagatctgacgaggaagcaggaagaaactgtagtccagataccg 441  
DB 3330 ttgaggggagcgagcagatctgacgaggaagcaggaagaaactgtagtccagataccg 3389  
QY 442 gccccaacacagcagcaacaagaatgagagagtttttggggagcagctggtatttcagacgt 501  
DB 3390 gccccaacacagcagcaacaagaatgagagagtttttggggagcagctggtatttcagacgt 3449  
QY 502 tggatcccgagggtttgacacttagcagcccaactacacccgttaacaaagaagaagg 561  
DB 3450 tggatcccgagggtttgacacttagcagcccaactacacccgttaacaaagaagaagg 3509  
QY 562 ggaattctctcgtgactctgagcagcagaaggaattgagtctatcaaaaagccctgctg 621  
DB 3510 ggaattctctcgtgactctgagcagcagaaggaattgagtctatcaaaaagccctgctg 3569  
QY 622 agcgcacctgctctgagccctccctgacgttaactaaacctttaccctttatgfgagag 681  
DB 3570 agcgcacctgctctgagccctccctgacgttaactaaacctttaccctttatgfgagag 3629  
QY 682 cgttaaggagtagcagcagagagttttaaccaaacccttagagacatgagagacactgtc 741  
DB 3630 cgttaaggagtagcagcagagagttttaaccaaacccttagagacatgagagacactgtc 3689  
QY 742 gctactactgtcaaaagagctgtgactcgtgacagtggttggccgtatgtctgaaggtc 801  
DB 3690 gctactactgtcaaaagagctgtgactcgtgacagtggttggccgtatgtctgaaggtc 3749  
QY 802 atcgcaagctgtgagcagatgagtaagaagcgtgagacaattgctttggagacgaatata 861  
DB 3750 atcgcaagctgtgagcagatgagtaagaagcgtgagacaattgctttggagacgaatata 3809  
QY 862 acttaataagcccccaatgcatgtgagacaatcgttcgagagcccccaagccagatgagt 921  
DB 3810 acttaataagcccccaatgcatgtgagacaatcgttcgagagcccccaagccagatgagt 3869

QY 922 accaagcccgcatgaccacatataaagcctgctcttcacagagaggtcattcgtc 981  
DB 3870 accaagcccgcatgaccacatataaagcctgctcttcacagagaggtcattcgtc 3929  
QY 982 ccacagccgctcttaacacctgccaactcttctgctgaagagactgtagaacctgact 1041  
DB 3930 ccacagccgctcttaacacctgccaactcttctgctgaagagactgtagaacctgact 3989  
QY 1042 catgattccatcaactttgatttagagagactgggtccgcaaggaacttaccagata 1101  
DB 3990 catgattccatcaactttgatttagagagactgggtccgcaaggaacttaccagata 4049  
QY 1102 ccgctgactggaagagtgctaacctggttcacatcgaggaagcagctatgtgtggaagtc 1161  
DB 4050 ccgctgactggaagagtgctaacctggttcacatcgaggaagcagctatgtgtggaagtc 4109  
QY 1162 aagaagatggtcgtgggagcgagctgtgtgagacggagccgcaagatcttgggacagacctg 1221  
DB 4110 aagaagatggtcgtgggagcgagctgtgtgagacggagccgcaagatcttgggacagacctg 4169  
QY 1222 ccggaaggaacttcagcgcaaaagctgagctcagtcagccctcaagacttgcgctg 1281  
DB 4170 ccggaaggaacttcagcgcaaaagctgagctcagtcagccctcaagacttgcgctg 4229  
QY 1282 gccgaagggaaatccataaacttatacggacagaggtatgaccttgcagctgcacac 1341  
DB 4230 gccgaagggaaatccataaacttatacggacagaggtatgaccttgcagctgcacac 4289  
QY 1342 gtacacggggccatctataaacaaggggttgccttacctagcagggaggaataaag 1401  
DB 4290 gtacacggggccatctataaacaaggggttgccttacctagcagggaggaataaag 4349  
QY 1402 aacaagaggaatcttaagccctatataaagccttatacttgcacaaaaggtcagatct 1461  
DB 4350 aacaagaggaatcttaagccctatataaagccttatacttgcacaaaaggtcagatct 4409  
QY 1462 atcacatgctctgtgacatcagaagaagcagaatctcatatctagaagggaacagatgct 1521  
DB 4410 atcacatgctctgtgacatcagaagaagcagaatctcatatctagaagggaacagatgct 4469  
QY 1522 gacccggttgcgaagcagcagccagcgtgttaacctcttcgacctatataagaaagcc 1581  
DB 4470 gacccggttgcgaagcagcagccagcgtgttaacctcttcgacctatataagaaagcc 4529  
QY 1582 aaagccccaagaaacccagacagacagatcacacctagaagacgtggcaagataaagaata 1641  
DB 4530 aaagccccaagaaacccagacagacagatcacacctagaagacgtggcaagataaagaata 4589  
QY 1642 gaccaggttctctgagactccggaggggacgtctatacctcatatggaagaaatccctg 1701  
DB 4590 gaccaggttctctgagactccggaggggacgtctatacctcatatggaagaaatccctg 4649  
QY 1702 ccccaaaaagaaaggtttagaataatgccaaacagatacatgcttcaaacccacttagaact 1761  
DB 4650 ccccaaaaagaaaggtttagaataatgccaaacagatacatgcttcaaacccacttagaact 4709  
QY 1762 aaacacctgagcagagttggtcagaacatccctcatcatgctctgaagctaccagaggt 1821  
DB 4710 aaacacctgagcagagttggtcagaacatccctcatcatgctctgaagctaccagaggt 4769  
QY 1822 gctgactcgtgtgtaaaatattgttgcctctcagagctggttaatgactaatccctcaga 1881  
DB 4770 gctgactcgtgtgtaaaatattgttgcctctcagagctggttaatgactaatccctcaga 4829  
QY 1882 ataacctccagaaagagactaaagggaagccacacagcgctcactatggaagtgacttc 1941  
DB 4830 ataacctccagaaagagactaaagggaagccacacagcgctcactatggaagtgacttc 4889  
QY 1942 actgaggtaaagccggtctaaatacaggaacaaatctatgtgttttgtgacacctt 2001  
DB 4890 actgaggtaaagccggtctaaatacaggaacaaatctatgtgttttgtgacacctt 4949  
QY 2002 tcagagtggttgagagcttactctactaagaaaggaacttcaacgggtgtgtgtaagaa 2061







Db 5250 ccccttgctctttagggtagaacaacccctggacagcttctggctgaccccttagaa 5309  
Qy 2362 ttgcttaacgggggaacccccccgcttgacagaattgcctcttcacatgtgtgtgtg 2421  
Db 5310 ttactctacgggggaacccccccatgtgtgaagaattgtctctgataatgtgtgacgtg 5369  
Qy 2422 ctgcttccacagccttctgtctcttagagctcaaggcgtcgaagtgtgtgacgagcg 2481  
Db 5370 ctgcttccacagccttctgtctcttagagctcaaggcgtcgaagtgtgtgacgagcg 5429  
Qy 2482 tggagcagctccgggagggcctactc---aggagagagcttgcgaattccacatgcctc 2538  
Db 5430 tggagcagctccgggagggcctactcagagagagagacttgcagatccacatgccttc 5489  
Qy 2539 caagttggagatcagctatgtttagacgcccacgctgacgaggaaccttgagagctgtgtg 2598  
Db 5490 caagttggagatcagctatgtttagacgcccacgctgacgaggaaccttcgagatcgtgtg 5549  
Qy 2599 aagggaacctatctcgtactcttgacacacacacagcgctgtgaaagtgcgaagaatcccc 2658  
Db 5550 aagggaacctatctcgtactcttgacacacacacagcgctgtgaaagtgcgaagaatcccc 5609  
Qy 2659 acctggtatcagatgcacatccacgcttaagccggcgccacccctccgagatcgaggctgaaagcc 2718  
Db 5610 acctggtatcagatgcacatccacgcttaagccggcgccacccctccgagatcgaggctgaaagcc 5669  
Qy 2719 gaaagaactgaaatcccttaagctcgcgccatccgctgctgttcccttaactctgaat 2778  
Db 5670 gaaagaactgaaatcccttaagctcgcgccatccgctgctgttcccttaactctgaat 5729  
Qy 2779 aactcctcaagtaatgtgtaaagcgtgtgtgacagcccgaaactccataaaccttalc 2838  
Db 5730 aactcctcagactaatgtgtatgtgcataagagacagcctgaactccataaaccttalc 5789  
Qy 2839 tctcactctgttactactagctacgctcgggtacaggtatataataacagactccaaggga 2898  
Db 5790 tctcactctgttactactagctacgctcgggtacaggtatataataacagactccaaggga 5849  
Qy 2899 ggcctccttggggagcctgtgtgcctgaatataatgtcctcctcagatcaatccctgg 2958  
Db 5850 ggcctccttggggagcctgtgtgcctgaatataatgtcctcctcagatcaatccctlag 5909  
Qy 2959 tctcaatgaacagccacaccccccagatgactccgtgcttaagggtttacgtttggccc 3018  
Db 5910 tct-----gacctcacccccagatctcctccatgctcagcagatttatgtttggccc 5960  
Qy 3019 aggaaccccaataatgaagaatattgtggaaatctcctaagattccttcttgaagaatg 3078  
Db 5961 aggaaccccaataatgaagaatattgtggaaatctcctaagattccttcttgaagaatg 6020  
Qy 3079 gacgtgacataacttctaatgaatgtggaatgtggaatgtgcagctcctcagcagaagaagt 3138  
Db 6021 gacgtgacataacttctaatgaatgtggaatgtggaatgtgcagcctcctcagcagaagaagt 6080  
Qy 3139 aagttactcttctgttaagaatccctacagatataataatataatgtgacatggagag 3198  
Db 6081 aagttactcttctgttaagaatccctacagatataataatataatgtgacatggagag 6140  
Qy 3199 atggaagaatgtgacagcggtacaaaagaatgtacgaataatgaagaatgaagt 3254  
Db 6141 tagaactgtgaagcccccagtgctctccttcagacctaatactaaataaagtt 6196

## RESULT 9

AAC67022  
ID AAC67022 standard; DNA: 4918 BP.

XX AAC67022;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 22.

XX XX  
KW Xenotransplantation; infectious agent; vaccine; ds.  
OS Porcine endogenous retrovirus.  
PN WO200071726-A1.  
PD 30-NOV-2000.  
PF 24-MAY-2000; 2000WO-US14296.  
PR 24-MAY-1999; 99US-0135631.  
PA (MAYO-) MAYO MEDICAL VENTURES.  
PI Federspiel MJ;  
DR WPI; 2001-032041/04.  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
PS Claim 16; Page 109-111; 144pp; English.  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.  
XX Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;  
SO

Query Match 82.0%; Score 2721.4; DB 22; Length 4918;  
Best Local Similarity 97.7%; Pred. No. 0;  
Matches 2771; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

Qy 486 tggatttggagactgtgagatccgggggttggagacttgcagacccacttaccgct 545  
Db 1 tggatttggagactgtgagatccgggggttggagacttgcagacccacttaccgct 60  
Qy 546 aaccaaagaagaagggaattctcctgtggtcctcgtgacacacagaaggcattgtatg 605  
Db 61 aaccaaagaagaagggaattctcctgtggtcctcgtgacacacagaaggcattgtatg 120  
Qy 606 caaaaggccctgtgagcgacactgtctgcccctcctgacgtlaactaaaccccttac 665  
Db 121 caaaaggccctgtgagcgacactgtctgcccctcctgacgtlaactaaaccccttac 180  
Qy 666 ccttatgtgagtagagtagagtagagtagagtagagtagagtagagtagagtagagtag 725  
Db 181 ccttatgtgagtagagtagagtagagtagagtagagtagagtagagtagagtagagtag 240  
Qy 726 atggaaggaagactgttgcctacactgtcaaaagactgtatccctgttaacagagtggtg 785  
Db 241 atggaaggaagactgttgcctacactgtcaaaagactgtatccctgttaacagagtggtg 300  
Qy 786 cgtatgtcctaaagactgtcagctgtgcccatactgtgtcaagaagcgtgaacaattgac 845  
Db 301 cgtatgtcctaaagactgtcagctgtgcccatactgtgtcaagaagcgtgaacaattgac 360  
Qy 846 ttggagacagaataactgttaataagccccccatgtctggagaacaatgttggagacc 905  
Db 361 ttggagacagaataactgttaataagccccccatgtctggagaacaatgttggagacc 420  
Qy 906 ccagacagatgtgatacgaacgcccgcagatgaacacatacaagaagcctgtcttcacaga 965  
Db 421 ccagacagatgtgatacgaacgcccgcagatgaacacatacaagaagcctgtcttcacaga 480  
Qy 966 gagggtacttctgcttcacacagcgtctcaacccctgcacactcttctgctgaagaagc 1025



This sequence represents the purified porcine retroviral cDNA sequence of Tsukuba-1 and contains the putative coding regions for viral proteins Gag, Pol and Env. This sequence and PCR fragments generated from the sequence (see AAT74812-T74882) could be used to screen organs for porcine retroviruses prior to orthotransplantation. Transplantation can increase the likelihood of retroviral activation if intact and infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantable donors. It can be used to detect mutations, genetic lesions or viral recombinants and to determine the histological localisation of activated retroviruses. Using Polymerase Chain Reaction DNA Quantitation (Pdq) on blood mononuclear cells, infectivity

QY	22	gagcgtctctctgcgcgtgagatataaccccactagacaaacactttgtcccttgaaatg	81
Db	5495	gatgcctctctctgcgcgtgagatataaccccactagacaaacactttgtcccttgaaatg	5554
QY	82	agagatccagtgatcagggaaagaaacccggcagctcaaccttgaccccgactgtcccaagggttc	141
Db	5555	agagatccagtgatcagggaaagaaacccggcagctcaaccttgaccccgactgtcccaagggttc	5614
QY	142	aaagaacccccgcgcacactctttgacgaagccctcaacagggaaacctgagccaaattcaagatc	201
Db	5615	aaagaacccccgcgcacactctttgacgaagccctcaacagggaaacctgagccaaattcaagatc	5674
QY	202	caaacacccctcaaggatgacccctcctccacgtacgtgatatccctgtctcgtcggggagccacc	261
Db	5675	caaacacccctcaaggatgacccctcctccacgtacgtgatatccctgtctcgtcggggagccacc	5734
QY	262	aaacacagactgctcttaagaagtgacgaagggcactactgctggaattgtctgaactaagctac	321
Db	5735	aaacacagactgctcttaagaagtgacgaagggcactactgctggaattgtctgaactaagctac	5794
QY	322	agagcctctcgtctaagaagggcccgagattttgacgagagagagtgtaacatacttggggttaagt	381
Db	5795	agagcctctcgtctaagaagggcccgagattttgacgagagagagtgtaacatacttggggttaagt	5854
QY	382	ttgcgggggcgggcagcagatgagctgacggagagcaggaagaaacctgtaagctccagataccg	441
Db	5855	ttgcgggggcgggcagcagatgagctgacggagagcaggaagaaacctgtaagctccagataccg	5914
QY	442	gcccacaacacaaacaaacaaagtagaagaagtttttggggcagaactgattttgcagactg	501
Db	5915	gcccacaacacacaaacaaacaaagtagaagaagtttttggggcagaactgattttgcagactg	5974
QY	502	tggatcccggggttttcgcgaacttagacagcccacacttaccgcgttaacccaagaagaaagg	561
Db	5975	tggatcccggggttttcgcgaacttagacagcccacacttaccgcgttaacccaagaagaaagg	6034
QY	562	ggattctcttggtcttccttgagcacacgaaggaatttgatgctatacaaaaggccctgtcg	621
Db	6035	ggattctcttggtcttccttgagcacacgaaggaatttgatgctatacaaaaggccctgtcg	6094
QY	622	agcgacacccgtctctggccctcccttgacgtaacgttaacaaaccttbaacctttagtggatga	681
Db	6095	agcgacacccgtctctggccctcccttgacgtaacgttaacaaaccttbaacctttagtggatga	6154
QY	682	cgtlaaaggagctagcccgagagagttttaaccccaaaccttagacacatgtagagagaaacctgtc	741
Db	6155	cgtlaaaggagctagcccgagagagttttaaccccaaaccttagacacatgtagagagaaacctgtc	6214
QY	742	gacctacctgtcaagaagcttgatcctctgttagccaagtggtgtggccgtaigtctgaaaggct	801
Db	6215	gacctacctgtcaagaagcttgatcctctgttagccaagtggtgtggccctatgctcgtgaaggct	6274
QY	802	atcgcagcgcgttgagccatactggttcaagggagcgtcgtaacaaattgaccttggagac- agatat	860
Db	6275	atcgcagcgcgttgagccatactggttcaagggagcgtcgtaacaaattgaccttggagagatat	6334
QY	861	aacctgtaatagcccccacatgcatctggagaacaactcgtttcgccagccccccagaccgatatg	920
Db	6335	aacctgtaatagcccccacatgcatctggagaacaactcgtttcgccagccccccagaccgatatg	6394
QY	921	gacccaagcccgcatgaccacattcaaaagcctgtcttccaaagagaggttcaacttcgc	980
Db			

Db 6395 gaccacgcccgcattgaccacatataaagcgtctctcaccagagaggtcaggttcgc 6454  
 Qy 981 tcacacacgcgcctcaaaccttcacactctctgcttgagagagagatgaaacacagtgac 1040  
 Db 6455 tcacacacgcgcctcaaaccttcacactctctgcttgagagagagatgaaacacagtgac 6514  
 Qy 1041 tcatattgcatcaactatgtattgagagagagagagagagagagagagagagagagagag 1100  
 Db 6515 tcatattgcatcaactatgtattgagagagagagagagagagagagagagagagagagag 6574  
 Qy 1101 accgcgtgactggaag 1160  
 Db 6575 accgcgtgactggaag 6634  
 Qy 1161 taagagagatgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1220  
 Db 6635 taagagagatgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 6694  
 Qy 1221 gccggaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 1280  
 Db 6695 gccggaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 6754  
 Qy 1281 gccggaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 1340  
 Db 6755 gccggaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 6814  
 Qy 1341 cgtacacggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1400  
 Db 6815 cgtacacggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 6874  
 Qy 1401 gaaacaaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 1460  
 Db 6875 gaaacaaaggaacttcagcgcaaaaggctgagctatggtcctcagcaagcttgcggtc 6934  
 Qy 1461 tatatacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1520  
 Db 6935 tatatacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 6994  
 Qy 1521 tgaccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1580  
 Db 6995 tgaccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 7054  
 Qy 1581 caaagcccccagacccagagacagatcacacccctagaagagagagagagagagagag 1640  
 Db 7055 caaagcccccagacccagagacagatcacacccctagaagagagagagagagagag 7114  
 Qy 1641 agaccagttctgagaccccgagagagagagagagagagagagagagagagagagagag 1700  
 Db 7115 agaccagttctgagaccccgagagagagagagagagagagagagagagagagagagag 7173  
 Qy 1701 gccccacaagaagggttagaataatgccaac--agataacatgctaccacactagaa 1759  
 Db 7174 gccccacaagaagggttagaataatgccaac--agataacatgctaccacactagaa 7233  
 Qy 1760 ctataacacgcgcagaggttggtcagagacatcccttataatgctcgaaggtcaccagag 1819  
 Db 7234 ctataacacgcgcagaggttggtcagagacatcccttataatgctcgaaggtcaccagag 7293  
 Qy 1820 tggctgactcgtggtcgaacatgtgtggtcgtcagagagagagagagagagagagagag 1879  
 Db 7294 tggctgactcgtggtcgaacatgtgtggtcgtcagagagagagagagagagagagagag 7353  
 Qy 1880 gaatacctccagaaag 1939  
 Db 7354 gaatgctccagaaag 7413  
 Qy 1940 tcaactggtgaaagccggtcgaataacggaacaaataatcattggtttttgtagaacct 1999  
 Db 7414 tcaactggtgaaagccggtcgaataacggaacaaataatcattggtttttgtagaacct 7473  
 Qy 2000 ttcaagaatggtgag 2059  
 Db 7474 ttcaagaatggtgag 7533

Qy 2060 aaatactgaggaatttttccaaagatttggaataacttaaggtlaataaggtgcacacatg 2119  
 Db 7534 aaatactgaggaatttttccaaagatttggaataacttaaggtlaataaggtgcacacatg 7593  
 Qy 2120 gtcacgttctggtggtcagagtaagtcagagagagagagagagagagagagagagagag 2179  
 Db 7594 gtcacgttctggtggtcagagtaagtcagagagagagagagagagagagagagagagag 7653  
 Qy 2180 aactgcatggtcag 2239  
 Db 7654 aactgcatggtcag 7713  
 Qy 2240 ttaag 2299  
 Db 7714 ttaag 7773  
 Qy 2300 tgccttctggttctttaggtgagagagagagagagagagagagagagagagagagagag 2359  
 Db 7774 tgccttctggttctttaggtgagagagagagagagagagagagagagagagagagagag 7833  
 Qy 2360 aatgctctacag 2419  
 Db 7834 aatgctctacag 7893  
 Qy 2420 tgccttctggttctttaggtgagagagagagagagagagagagagagagagagagagag 2479  
 Db 7894 tgccttctggttctttaggtgagagagagagagagagagagagagagagagagagagag 7953  
 Qy 2480 cgtggaag 2536  
 Db 7954 cgtggaag 8013  
 Qy 2537 tccaaggtgag 2593  
 Db 8014 tccaaggtgag 8060

RESULT 11  
 AAF77725  
 ID AAF77725 standard; cDNA; 8060 BP.  
 XX  
 AC AAF77725;  
 XX  
 DT 23-MAY-2001 (first entry)  
 XX  
 DE Tsukuba-1 cDNA.  
 XX  
 KW Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.  
 XX  
 OS Porcine retrovirus.  
 XX  
 PN US6190861-B1.  
 XX  
 PD 20-FEB-2001.  
 XX  
 PF 13-DEC-1996; 96US-0766528.  
 XX  
 PR 14-DEC-1995; 95US-0572645.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Fishman JA;  
 XX  
 DR WPI: 2001-256211/26.  
 XX  
 PT Assessing risk of endogenous retroviruses in clinical practice and in  
 PT xenotransplantation, comprises using probe sequences derived from swine  
 PT or miniature swine retroviral genome -  
 XX  
 PS Claim 1; Fig 1; 127pp; English.  
 XX  
 CC The present invention relates to a method for screening a cell or tissue

CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g., the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match 72.9%; Score 2420.6; DB 22; Length 8060;  
 Best Local Similarity 97.5%; Pred. No. 0;  
 Matches 2502; Conservative 0; Mismatches 59; Indels 6; Gaps 4;

QY 22 gatgctctctctgctgagatlaaacccactagccaaacttttgctctcgatg 81  
 Db 5495 gatgctctctctgctgagatlaaacccactagccaaacttttgctctcgatg 5554  
 QY 82 agaatccaggtacgggaagaacgagcgagctacccctggaccgagcccaagggttc 141  
 Db 5555 agaatccaggtacgggaagaacgagcgagctacccctggaccgagcccaagggttc 5614  
 QY 142 aagaactcccgaccatcttgcagaaagccctacacagagacctggtcccaacttcagatc 201  
 Db 5615 aagaactcccgaccatcttgcagaaagccctacacagagacctggtcccaacttcagatc 5674  
 QY 202 caacacccctcagctgacccctcctcagctgagatgactgctctgctggagagccacc 261  
 Db 5675 caacacccctcagctgacccctcctcagctgagatgactgctctgctggagagccacc 5734  
 QY 262 aaacagagactgcttagaaggtacgaagcgacactgctggaattgtctgacaggtac 321  
 Db 5735 aaacagagactgcttagaaggtacgaagcgacactgctggaattgtctgacaggtac 5794  
 QY 322 agagcctctgcttaagaagcgccagatttgcagagagaggtlaacatactggtggtacagt 381  
 Db 5795 agagcctctgcttaagaagcgccagatttgcagagagaggtlaacatactggtggtacagt 5854  
 QY 382 ttgggggagcgagcgagctgctgagcgagcgagcgagaaacactgtatgccagataccg 441  
 Db 5855 ttgggggagcgagcgagctgctgagcgagcgagcgagaaacactgtatgccagataccg 5914  
 QY 442 gcccaacccacagccaacaagtgagagagtttttggggacagctggaatttgcagactg 501  
 Db 5915 gcccaacccacagccaacaagtgagagagtttttggggacagctggaatttgcagactg 5974  
 QY 502 tggatcccggttttgacacttagcagcccaactacccggttaacccaagaagaagag 561  
 Db 5975 tggatcccggttttgacacttagcagcccaactacccggttaacccaagaagaagag 6034  
 QY 562 ggaattccctgggctctgagcaaccagaagcaattgtagtctacaanaagcgctctg 621  
 Db 6035 ggaattccctgggctctgagcaaccagaagcaattgtagtctacaanaagcgctctg 6094  
 QY 622 agcgacactgctgagccctccctgagctaaactaaaccccttaaccccttatgtgagtg 681  
 Db 6095 agcgacactgctgagccctccctgagctaaactaaaccccttaaccccttatgtgagtg 6154  
 QY 682 cgtaaaggaatagcccgagaggttttaacccaacccctagagacatgtagagagactgct 741  
 Db 6155 cgtaaaggaatagcccgagaggttttaacccaacccctagagacatgtagagagactgct 6214  
 QY 742 gctactctgcaagaagacttgatctctgtagcaggtgtgtgcccgtatgtctgaagct 801  
 Db 6215 gctactctgcaagaagacttgatctctgtagcaggtgtgtgcccgtatgtctgaagct 6274  
 QY 802 atcgcaactgtgacatactgctcaaggaagctgacaaattgactttggagc-agaatat 860  
 Db 6275 atcgcaactgtgacatactgctcaaggaagctgacaaattgactttggagc-agaatat 6334

QY 861 aactgtaataagcccccatgcatctgagaaacatcgttgcgagccccagaccagatgat 920  
 Db 6335 aactgtaataagcccccatgcatctgagaaacatcgttgcgagccccagaccagatgat 6394  
 QY 921 gaccacagcccgatgacccactatcaaacgctgctctcacaagagaggtcactctgc 980  
 Db 6395 gaccacagcccgatgacccactatcaaacgctgctctcacaagagaggtcactctgc 6454  
 QY 981 tcacacagcgctctcaaaccttgccactcttctgctggaagagactgataaccagtgac 1040  
 Db 6455 tcacacagcgctctcaaaccttgccactcttctgctggaagagactgataaccagtgac 6514  
 QY 1041 tcagtgtgcacatcaactttgattagagagactgggtgcgcaagagacttaccagat 1100  
 Db 6515 tcagtgtgcacatcaactttgattagagagactgggtgcgcaagagacttaccagat 6574  
 QY 1101 accgctgactggaaggtgtaacctgtgtcactgacgggaagcagatgtgtggaag 1160  
 Db 6575 accgctgactggaaggtgtaacctgtgtcactgacgggaagcagatgtgtggaag 6634  
 QY 1161 taagaagatggtctggggtcgagctgtgaaacgggacccgacagatctggccagagcct 1220  
 Db 6635 taagaagatggtctggggtcgagctgtgaaacgggacccgacagatctggccagagcct 6694  
 QY 1221 gccggaaggaacttcagcgcaaaagctgagctcatggtccctcacgagcttttgagct 1280  
 Db 6695 gccggaaggaacttcagcgcaaaagctgagctcatggtccctcacgagcttttgagct 6754  
 QY 1281 ggcgaagggaaatccataaacaattatagcgacagcaggtatgtccttgcagctgaca 1340  
 Db 6755 ggcgaagggaaatccataaacaattatagcgacagcaggtatgtccttgcagctgaca 6814  
 QY 1341 cgtacacggggccatctataaacaaggggtgttaccttaagcagaggaaggaataa 1400  
 Db 6815 cgtacacggggccatctataaacaaggggtgttaccttaagcagaggaaggaataa 6874  
 QY 1401 gaacaagaggaatctcagcctatlaagaaccttlaaccttgcgaanaaagctgactat 1460  
 Db 6875 gaacaagaggaatctcagcctatlaagaaccttlaaccttgcgaanaaagctgactat 6934  
 QY 1461 tatacactgctctgagacatcagaagccaagatcatalactagaaggaacagatgac 1520  
 Db 6935 tatacactgctctgagacatcagaagccaagatcatalactagaaggaacagatgac 6994  
 QY 1521 tgaccgggttgcgaagcgagcgaccaggtgttaacttctgctctataatagaagcc 1580  
 Db 6995 tgaccgggttgcgaagcgagcgaccaggtgttaacttctgctctataatagaagcc 7054  
 QY 1581 caaagccccaagacccaagacagctacacccctagaagaactggaagagataaagaat 1640  
 Db 7055 caaagccccaagacccaagacagctacacccctagaagaactggaagagataaagaat 7114  
 QY 1641 agacagatctctgagactcggagggagcctgtcatalactcatalatgggaaggaatcct 1700  
 Db 7115 agacca-ttctctgagactcggagggagcctgtcatalactcatalatgggaaggaatcct 7173  
 QY 1701 gccccaacaagaaggttagaataatgtccaac-agaatacctgcttaacccaactaggaa 1759  
 Db 7174 gccccaacaagaaggttagaataatgtccacaagaatacctgcttaacccaactaggaa 7233  
 QY 1760 ctaaacacactgcaagcaatgtgtcagaacaatcccttcaatgttctgagctacagagag 1819  
 Db 7234 ctaaacacactgcaagcaatgtgtcagaacaatcccttcaatgttctgagctacagagag 7293  
 QY 1820 tggctgactcgggtgtaacaacatltgtgccctgcagcgtgttaatgtctaactcca 1879  
 Db 7294 tggctgactcgggtgtaacaacatltgtgccctgcagcgtgttaatgtctaactcca 7353  
 QY 1880 gaatacctcagaaggaagagactaaagggaagccacccagcgctcactgggaatgtgact 1939  
 Db 7354 gaatgctccacaggaagagactaaagggaagccacccagcgctcactgggaatgtgact 7413



QY 1940 tcactgagtaaaacgcgctaataacgaaacaatactatgtgtttttagaacact 1999  
 Db 7414 tcactgagtaaaacgcgctaataatgaaacaatactatgtgtttttagaacact 7473  
 QY 2000 tttaagatggtgtagagcttactactactaagaagaagacttaaccgtgtgtctaaga 2059  
 Db 7474 tttaagatggtgtagagcttactactactaagaagaagacttaaccgtgtgtctaaga 7533  
 QY 2060 aaatacggaggaatatttccaagatttgaatacctaagtaagagtcagacaatg 2119  
 Db 7534 aaatacggaggaatatttccaagatttgaatacctaagtaagagtcagacaatg 7593  
 QY 2120 gtccagcttcgttgcacgaagtaagtaaggaagactgcccagaatattggtgatttga 2179  
 Db 7594 gtccagcttcgttgcacgaagtaagtaaggaagactgcccagaatattggtgatttga 7653  
 QY 2180 aactgacttgcatacagaccccaagctcaggaagctgtagagagatgataagaacca 2239  
 Db 7654 aactgacttgcatacagaccccaagctcaggaagctgtagagagatgataagaacca 7713  
 QY 2240 ttaaagagaccccttaaccaatgacacacagagactgcatatgattgtgactctcc 2299  
 Db 7714 ttaaagagaccccttaaccaatgacacacagagactgcatatgattgtgactctcc 7773  
 QY 2300 tgcctcttgcgttttaaggtgtagaagaacaccccttggaagcttgggtgtgacccctatg 2359  
 Db 7774 tgcctcttgcgttttaaggtgtagaagaacaccccttggaagcttgggtgtgacccctatg 7833  
 QY 2360 aatgactcagggagagaccccccgcgttgcagaaatgtgcttgcatacagtagtgcgatg 2419  
 Db 7834 aatgactcagggagagaccccccgcgttgcagaaatgtgcttgcatacagtagtgcgatg 7893  
 QY 2420 tgcgtcttcccaagccttctctctagagctcaagcgctcgatgtggtgtagagcaagcag 2479  
 Db 7894 tgcgtcttcccaagccttctctctagagctcaagcgctcgatgtggtgtagagcaagcag 7953  
 QY 2480 cgtggaagcagctcccggaagccttactc---aaggaagagagcttgcacatcgcgt 2536  
 Db 7954 cgtggaagcagctcccggaagccttactc---aaggaagagagcttgcacatcgcgt 2603  
 QY 2537 tccaagctgagatcagctatgtagagcagcagctgacaggaac 2583  
 Db 8014 tccaagctgagatcagctatgtagagcagcagctgacaggaac 8060

RESULT 12  
 AAC67021  
 ID AAC67021 standard; DNA; 6076 BP.  
 AC AAC67021.  
 AC AAC67021.  
 DT 27-MAR-2001 (first entry)  
 DE PERV env protein coding sequence SEQ ID NO: 21.  
 KM Xenotransplantation; infectious agent; vaccine; ds.  
 OS Porcine endogenous retrovirus.  
 PN WO200071726-A1.  
 PD 30-NOV-2000.  
 PF 24-MAY-2000; 2000MO-US14296.  
 PR 24-MAY-1999; 990S-0135631.  
 PA (MAYO-) MAYO MEDICAL VENTURES.  
 PI Federspiel MJ;  
 DR WPI; 2001-032041/04.  
 XX

PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 PS Claim 16; Page 107-109; 144pp; English.  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 CC XX  
 SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match 72.8%; Score 2415.6; DB 22; Length 6076;  
 Best Local Similarity 98.0%; Pred. No. 0;  
 Matches 2445; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 22 gatgcctctctgcctgagattacaccccaacacacacttttgcctcgaatg 81  
 Db 3568 gatgcctctctgcctgagattacaccccaacacacacttttgcctcgaatg 3627  
 QY 82 agagatcgaagtagcgggaagaacgcgagctacactgaccgcagctcccaaggttc 141  
 Db 3628 agagatcgaagtagcgggaagaacgcgagctacactgaccgcagctcccaaggttc 3687  
 QY 142 aagaactcccccacacatcttgaagaaagccctacaaagagccctgccaactttagatc 201  
 Db 3688 aagaactcccccacacatcttgaagaaagccctacaaagagccctgccaactttagatc 3747  
 QY 202 caacacccctcaagtgaccctctccagtaagtgatgacctgtcttgcggagacacc 261  
 Db 3748 caacacccctcaagtgaccctctccagtaagtgatgacctgtcttgcggagacacc 3807  
 QY 262 aaacagagactgtttagaagtagcgaagcactacgctggaattgtcgaactgtgctac 321  
 Db 3808 aaacagagactgtttagaagtagcgaagcactacgctggaattgtcgaactgtgctac 3867  
 QY 322 aagactctctgaagaaagcccaagatttgcagagagaggttaacatacttgggttacaqt 381  
 Db 3868 aagactctctgaagaaagcccaagatttgcagagagaggttaacatacttgggttacaqt 3927  
 QY 382 ttgcgggagcggagcagatgctgtagcaggaagcaggaagaaactgtagctccagatccg 441  
 Db 3928 ttgcgggagcggagcagatgctgtagcaggaagcaggaagaaactgtagctccagatccg 3987  
 QY 442 gccccaacacagacaaagaaagtagaagattttgggagagagctggaatttgcagactg 501  
 Db 3988 gccccaacacagacaaagaaagtagaagattttgggagagagctggaatttgcagactg 4047  
 QY 502 tgaatcccgagggttgcgacacttgaacgcccacttaccgcttaaccacaaagaaaggg 561  
 Db 4048 tgaatcccgagggttgcgacacttgaacgcccacttaccgcttaaccacaaagaaaggg 4107  
 QY 562 ggaattctctgggctcctgagcaccagaagaagcatttgaatgctatcaaaaagccctgctg 621  
 Db 4108 ggaattctctgggctcctgagcaccagaagaagcatttgaatgctatcaaaaagccctgctg 4167  
 QY 622 agcgcaactgctctgcccctccctgaagtaactaaaccccttaacccttaagttagag 681  
 Db 4168 agcgcaactgctctgcccctccctgaagtaactaaaccccttaacccttaagttagag 4227  
 QY 682 cgtlaagggagtagcccgagagattttaaccacaaaccccttagacatgtagagagactgtt 741  
 Db 4228 cgtlaagggagtagcccgagagattttaaccacaaaccccttagacatgtagagagactgtt 4287  
 QY 742 gctactcgttcaaaaagacttgatcctgtagcaggtgtgtgcccgtatgtctgaagct 801  
 Db 4288 gctactcgttcaaaaagacttgatcctgtagcaggtgtgtgcccgtatgtctgaagct 4347  
 QY 802 atcgagcgtgtgcatatcctgttcaagagcgtgacaacattgacttgggacagaata 861





PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
XX  
XX  
PS Claim 16; Page 117-119, 144pp; English.

XX  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during recombination. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC CC  
XX CC transplant recipient.

Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

Query Match 72.7%; Score 2414; DB 22; Length 6076;  
Best Local Similarity 98.0%; Pred. No. 0; Mismatches 50; Indels 0; Gaps 0;  
Matches 2444; Conservative 0;

QY 22 gatgcctctctgctgagattacacccactagccacacattttgctcgaatg 81  
DB 3568 gatgcctctctgctgagattacacccactagccacacattttgctcgaatg 3627  
QY 82 agagatccaggtacgggaagaacccgggagctacactgagcccgactgcccagaaggttc 141  
DB 3628 agagatccaggtacgggaagaacccgggagctacactgagcccgactgcccagaaggttc 3687  
QY 142 aagaactcccgacactcttgacgaagccctacacagagaccctggccaaactcagatc 201  
DB 3688 aagaactcccgacactcttgacgaagccctacacagagaccctggccaaactcagatc 3747  
QY 202 caacacccctcaggtgacccctccacgtacgtgagatgacctgctctgagggagccacc 261  
DB 3748 caacacccctcaggtgacccctccacgtacgtgagatgacctgctctgagggagccacc 3807  
QY 262 aaaaagagagcttagaagtaagaagagctactgctggaattgtctgacagatc 321  
DB 3808 aaaaagagagcttagaagtaagaagagctactgctggaattgtctgacagatc 3867  
QY 322 agagccctctgctgaagaagccagatttgagagagagtaacatacttgggggtacagt 381  
DB 3868 agagccctctgctgaagaagccagatttgagagagagtaacatacttgggggtacagt 3927  
QY 382 ttgcgggagcgagcagatgagtcgaggaagcaggaagaacatgtatgccagatacgg 441  
DB 3928 ttgcgggagcgagcagatgagtcgaggaagcaggaagaacatgtatgccagatacgg 3987  
QY 442 gcccacacacagcacaacaaagtgagagagtttttgggagacgtgatttcagactc 501  
DB 3988 gcccacacacagcacaacaaagtgagagagtttttgggagacgtgatttcagactc 4047  
QY 502 tggatcccgagggttgagacatttagaccccaactacacggctaaacaaagaagaagg 561  
DB 4048 tggatcccgagggttgagacatttagaccccaactacacggctaaacaaagaagaagg 4107  
QY 562 ggaattctctggtgctctgagcagcagagaagcatttgaatcctcaaaagccctgctg 621  
DB 4108 ggaattctctggtgctctgagcagcagagaagcatttgaatcctcaaaagccctgctg 4167  
QY 622 aggcacactgctgagccctccctgagcttaactaaacctttacccttttggatgag 681  
DB 4168 aggcacactgctgagccctccctgagcttaactaaacctttacccttttggatgag 4227  
QY 682 cgtaaaggagtagcccgagagattttaaccaaacccttagagacatgagagagacgttc 741  
DB 4228 cgtaaaggagtagcccgagagattttaaccaaacccttagagacatgagagagacgttc 4287  
QY 742 gctaacctgtcaaaagagcttgatcctgtagccagtggttggccgtatgtctgaagct 801  
DB 4288 gctaacctgtcaaaagagcttgatcctgtagccagtggttggccgtatgtctgaagct 4347  
QY 802 atcgacagctgtgacacatcgttgaagaagcgtgacaaattgacttgggagacgaatata 861  
DB 4348 atcgacagctgtgacacatcgttgaagaagcgtgacaaattgacttgggagacgaatata 4407

QY 862 actgtaataagcccccatgcatctgagagaacatcgttcggacgccccagacagatgagatg 921  
DB 4408 actgtaataagcccccatgcatctgagagaacatcgttcggacgccccagacagatgagatg 4467  
QY 922 accaagcgccagatgacccacatcaaaagcgtgtctctacagagaaggtcattctgct 981  
DB 4468 accaagcgccagatgacccacatcaaaagcgtgtctctacagagaaggtcattctgct 4527  
QY 982 ccacacagcagctctcaacccctgacactcttctgcttaagaagagatgataacacgtgact 1041  
DB 4528 ccacacagcagctctcaacccctgacactcttctgcttaagaagagatgataacacgtgact 4587  
QY 1042 catgattgccatcaactattgattgagagagactgggttcgcaagaagccttacagacata 1101  
DB 4588 catgattgccatcaactattgattgagagagactgggttcgcaagaagccttacagacata 4647  
QY 1102 ccgctgacttgagaagagctgtaactggtctactgaacgaacacatgctgtgtgaaggt 1161  
DB 4648 ccgctgacttgagaagagctgtaactggtctactgaacgaacacatgctgtgtgaaggt 4707  
QY 1162 aagaagatggtctggtggcgagctggttgagacggagcccgacagatctggccagacgctg 1221  
DB 4708 aagaagatggtctggtggcgagctggttgagacggagcccgacagatctggccagacgctg 4767  
QY 1222 ccggaaggaactctcagcgcaaaagcttgagctcaltgagccctacagcaagctttgcgctg 1281  
DB 4768 ccggaaggaactctcagcgcaaaagcttgagctcaltgagccctacagcaagctttgcgctg 4827  
QY 1282 gccgaaggaataatccataacattatagacagcagatgtgcttggagagcgccac 1341  
DB 4828 gccgaaggaataatccataacattatagacagcagatgtgcttggagagcgccac 4887  
QY 1342 gtaacaggggcccataataaacaaggggttgcttactcctaagcagggaggaataaag 1401  
DB 4888 gtaacaggggcccataataaacaaggggttgcttactcctaagcagggaggaataaag 4947  
QY 1402 aaaaaagagaaattcctaagcctatagagccttaacttggccaaaaggtcagctat 1461  
DB 4948 aaaaaagagaaattcctaagcctatagagccttaacttggccaaaaggtcagctat 5007  
QY 1462 atacactgctctgagacatcagaagaaccaaagatccatcatcttaagaaggagacagatgct 1521  
DB 5008 atacactgctctgagacatcagaagaaccaaagatccatcatcttaagaaggagacagatgct 5067  
QY 1522 gacccggttgcacaaagcagcagccagcagctgttlaaccttctgctcataatagaaagccc 1581  
DB 5068 gacccggttgcacaaagcagcagccagcagctgttlaaccttctgctcataatagaaagccc 5127  
QY 1582 aaagccccagaacccaagacagatgacaccccttagaagactggaagagataaaagaata 1641  
DB 5128 aaagccccagaacccaagacagatgacaccccttagaagactggaagagataaaagaata 5187  
QY 1642 gaccaggttcttgagacatccggaagggagcctgtctactactatgtgggaagaaatccctg 1701  
DB 5188 gaccaggttcttgagacatccggaagggagcctgtctactactatgtgggaagaaatccctg 5247  
QY 1702 ccccaaaagaaggtttagaataatgttcaacagatatacgttctaaccacacttagaact 1761  
DB 5248 ccccaaaagaaggtttagaataatgttcaacagatatacgttctaaccacacttagaact 5307  
QY 1762 aaacactgcaagcagttggttcagaacatctccttcaatgtcttgaagctacagagatg 1821  
DB 5308 aaacactgcaagcagttggttcagaacatctccttcaatgtcttgaagctacagagatg 5367  
QY 1822 gctgactcgtgtgtaaaatgtgtgcccgtgacggctggttaagtcaatccctccaga 1881  
DB 5368 gctgactcgtgtgtaaaatgtgtgcccgtgacggctggttaagtcaatccctccaga 5427  
QY 1882 ataacctcaggaagaagactaagggaagccacccagcgctcactctggaaggtgacttc 1941  
DB 5428 ataacctcaggaagaagactaagggaagccacccagcgctcactctggaaggtgacttc 5487



Oy	321t	gcaacagcgggtccaaaaagaattaccgaataaagcaataaagcttcattcgtagacct	3270'
Db	5285	gcacaagcgggtccaaaaagaatgacgaataaagcaataaagcttcattcgtagacct	5344'
Oy	3271	agataacttaaaaaaaagtgttcacttaaaaaaaaaaaaaaaaaaaaaa	3317
Db	5345	agataacttaaaaaaaagtgttcacttgtaaaaaggaaacaagaanaata	5391
<hr/>			
RESULT#	15		
ID	AAT74883		
	AAT74883	standard; cDNA; 7393 BP.	
XX			
AC	AAT74883;		
XX			
DT	09-FEB-1998	(first entry)	
XX			
DE	Porcine retrovirus cDNA (defective).		
XX			
KW	Retrovirus; porcine; GAG protein; POL protein; ENV protein;		
KM	xenotransplantation; infectious; provirus; organ transplant; donor;		
XX	activated virus; PCR; ss.		
OS	Porcine retrovirus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	598..2172	
		/+tag= a	
FT	mat_peptide	598..2169	
FT		/+tag= b	
FT		/note= "putative GAG protein"	
FT	CDS	2320..4737	
FT		/+tag= c	
FT		/note= "putative POL coding region (partial) as described in the specification"	
FT	mat_peptide	2320..3522	
FT		/+tag= d	
FT		/note= "putative POL protein (partial)"	
FT	mat_peptide	3516..4328	
FT		+tag= e	
FT		/note= "putative POL protein (partial)"	
FT	mat_peptide	4332..4748	
FT		/+tag= f	
FT	CDS	4738..6725	
FT		/+tag= g	
FT		/note= "putative ENV coding region (partial) as described in the specification"	
FT	mat_peptide	4752..6722	
FT		/+tag= h	
FT		/note= "ENV protein (partial)"	
PN	WO9721836-A1.		
XX			
PD	19-JUN-1997.		
XX			
PF	13-DEC-1996;	96WO-US19660.	
XX			
PR	14-DEC-1995;	95US-0572645.	
XX			
PA	(GEHO ) GEN HOSPITAL CORP.		
XX			
PI	Fishman JA;		
XX			
DR	WPI; 1997-332804/30.		
DR	P-PSDB; AAM32091-W32095.		
XX			
PT	New nucleic acid from porcine retro.viruses - used for detecting		
PT	viruses in transplant or other tissue and for assessing risk of		
PT	transmitting infection to graft recipient		
PS	Claim 16; Fig 2; 128bp; English.		
XX			

Query Match	53.4%;	Score 1771.8;	DB 18;	Length 7393;
Best Local Similarity	95.1%;	Pred. No. 0;		
Matches 1871; Conservative	0;	Mismatches 17;	Indels 79;	Gaps 11;

QY	2071	gaaatcttccagaatttgnaatacctaaggaataaggttcagacaatcgtccagcttc	2130
Db	4224	gaaatcttccagaatttgnaatacctaaggaataaggttcagacaatcgtccagcttc	4283
QY	2131	gttcgccaagttaagtcagaggaagctggccaagatatctggagatctgataactgcattgt	2190
Db	4284	gttcgccaagttaagtcagaggaagctggccaagatatctggagatctgataactgcattgt	4343
QY	2191	gatacagaagcccaagaagctcagacagatgagagaggaatgaaatagacaactaaagaagcc	2250
Db	4344	gatacagaagcccaagaagctcagacagatgagagaggaatgaaatagacaactaaagaagcc	4403
QY	2251	cttaccaaatctgacacagagagctggcattaaatgatatgagatgctctcgtccctctgtg	2310
Db	4404	cttaccaaatctgacacagagagctggcattaaatgatatgagatgctctcgtccctctgtg	4463
QY	2311	cttttaaggttgagaaacaaccccttgacaagtttgagctggacccctatgatatctctaac	2370
Db	4464	cttttaaggttgagaaacaaccccttgacaagtttgagctggacccctatgatatctctaac	4523
QY	2371	ggggagaccccccccttgagcaaaaattgcctttgacacatagtgacatgctgcctttcc	2430
Db	4524	ggggagaccccccccttgagcaaaaattgcctttgacacatagtgacatgctgcctttcc	4583
QY	2431	cagcctctgtctctagctcaaggcgtctcagatggtgtgagagcagcgatggaagcag	2490
Db	4584	cagcctctgtctctagctcaaggcgtctcagatggtgtgagagcagcgatggaagcag	4643
QY	2491	ctccggaggggacctactcagggaggaacttgcaagttccacactgcctccaaagtggaaat	2550
Db	4644	ctccggaggggacctactcagggaggaacttgcaagttccacactgcctccaaagtggaaat	4703
QY	2551	tcaagctatgtttagacgcacaccgttgacagaaaactcgcagactcgggtggaagggaaactat	2610
Db	4704	tcaagctatgtttagacgcacaccgttgacagaaaactcgcagactcgggtggaagggaaactat	4763
QY	2611	ctcgtacattttgacacacccaagcgcgtgtgaaagtctgaaggaatccccaactgtgacat	2670
Db	4764	ctcgtacattttgacacacccaagcgcgtgtgaaagtctgaaggaag-----	4804
QY	2671	gcataccacgttaagcggcgccacctcccgatccgggttggaagccgaaagaatgaa	2730
Db	4805	-----	4804
QY	2731	aatcccttaagctgcgcctccatcgcgttggtctccttaactgttcaataactccctaagt	2790
Db	4805	aatcccttaagctgcgcctccatcgcgttggtctccttaactgttcaataactccctaagt	4864
QY	2791	taatgtgtaaaagcctctgtggagacgcccgaactcccaataaaccttatctctcaactggtc	2850
Db	4865	taatgtgtaaaagcctctgtggagacgcccgaactcccaataaaccttatctctcaactggtc	4924
QY	2851	acttaactaacccgttaacagtatataatctaaacgcactcaaggggaggtccctctggg	2910
Db	4925	acttaactaacccgttaacagtatataatctaaacgcactcaaggggaggtccctctggg	4984
QY	2911	gaactgtgtgcctgaaatatatgtctgcctccatcagtaaatccctggtctcaatgacca	2970
Db	4985	gaactgtgtgcctgaaatatatgtctgcctccatcagtaaatccctggtctcaatgacca	5044
QY	2971	ggccacaccccccgatgactccggtgtctaaaggttttaagctttggccaagacccccaaa	3030
Db	5045	ggccacaccccccgatgactccggtgtctaaaggttttaagctttggccaagacccccaaa	5104
QY	3031	taatgaagaataatctgtggaatactccagagattcctttgcaagaatctgagctgataac	3090
Db	5105	taatgaagaataatctgtggaatactccagagattcctttgcaagaatctgagctgataac	5164
QY	3091	ttctaaatgattgggaattggaaatggccaagttcttcagcaagaacagatlaagttaactctt	3150
Db	5165	ttctaaatgattgggaattggaaatggccaagttcttcagcaagaacagatlaagttaactctt	5224

```

QY 3151 tgltaacaatcctaccagttataatcaatttaattatgccaatggagatggaagaattg 3210
    |||||||
Db 5225 tgltaacaatcctaccagttataatcaatttaattatgccaatggagatggaagaattg 5284
    |||||||
QY 3211 gcaacagcgggtacaaaaagatgtacgaataaagcaataaagctgtcattcgttagaact 3270
    |||||||
Db 5285 gcaacagcgggtacaaaaagatgtacgaataaagcaataaagctgtcattcgttagaact 5344
    |||||||
QY 3271 agattacttaaaaataagtttcaactaaaaaagaaaaaagaaaaa 3317
    |||||||
Db 5345 agattacttaaaaataagtttcaactgaaaaaagaaaaaagaaaaa 5391
    |||||||

```

Search completed: February 24, 2002, 02:48:16  
 Job time: 18462 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 02:48:16 ; Search time 624.99 Seconds

(without alignments) 11260.633 Million cell updates/sec

Title: US-09-171-553B-3

Perfect score: 8209  
Sequence: 1 gtgtgttcgcgccttgggccc.....aaaaaaaaaaaaaaaaaaaaa 8209

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N\_Geneseq\_1101:\*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8209	100.0	8209	18 AAV09700	Porcine retrovirus
2	8171.8	99.5	8196	18 AAV09699	Porcine retrovirus
3	6282.4	76.5	8132	22 AAT7727	Nucleotide sequenc
4	6141	74.8	7892	18 AAT74884	Miniature swine re
5	5637.2	68.7	7873	22 AAC67023	PERV env protein c
6	5528.2	67.3	7362	22 AAC67019	PERV env protein c
7	5489.8	66.9	7333	22 AAT7726	Defective retrovir
8	5416.6	66.0	7393	18 AAT74883	Porcine retrovirus
9	5025.8	61.2	6076	22 AAC67021	PERV env protein c
10	5025.8	61.2	6076	22 AAC67032	PERV env protein c
11	4986	60.7	8060	18 AAT74811	Porcine retrovirus

12	4986	60.7	8060	22 AAT77725	Tsukuba-1 cDNA.
13	3755.4	45.7	4402	22 AAC67020	PERV env protein c
14	3376.6	41.1	3482	22 AAV82749	Pig endogenous ret
15	3295.2	40.1	4918	22 AAC67022	PERV env protein c
16	32916	35.5	3320	20 AAV09698	Porcine retrovirus
17	2295.8	28.0	8655	20 AAV69750	Mus dunni endogeno
18	2244.4	27.3	8088	21 AA245540	Complete nucleotid
19	2234.4	27.2	8535	15 AA073731	Galy SEAFO genome.
20	1941.6	23.7	1974	18 AAV09703	Porcine retrovirus
21	1747	21.3	8323	16 AAO81042	Friend murine leuk
22	1747	21.3	8323	18 AAT80058	LTR clone of FB29
23	1747	21.3	8323	18 AAT07063	Friend ecotropic m
24	1737.4	21.2	8202	17 AAT09280	Novel AMP/MCF Viru
25	1735.6	21.1	8323	15 AA058700	LTR clone of FB29
26	1700	20.7	8278	22 AAC89138	Canine retrovirus
27	1697.8	20.7	8440	17 AAT13265	Pely F6a provirus
28	1696.6	20.7	8892	17 AAO94266	Osteoinductive ret
29	1687.4	20.6	8332	18 AAV05061	Moloney murine leu
30	1687.4	20.6	8332	18 AAV42715	Moloney murine leu
31	1687.4	20.6	8332	19 AAV19551	Retroviral vector
32	1687.4	20.6	8332	19 AAV15346	Retroviral gag gen
33	1687.4	20.6	8332	20 AAX28143	MLV gag, gag-pol
34	1687.4	20.6	8332	21 AAA40297	MLV gag, gag-pol
35	1687.4	20.6	8332	21 AA290064	Moloney murine leu
36	1687.4	20.6	8332	21 AA245539	Complete nucleotid
37	1686	20.5	8332	16 AAT08452	MoMLV gag/pol gene
38	1684.2	20.5	8332	17 AAT07374	MoMLV gag/pol gene
39	1680.4	20.5	8330	20 AAX77178	MoMLV gene sequenc
40	1596	19.4	10367	18 AAT80057	Sequence of plasmid
41	1596	19.4	10367	20 AA207067	Glycopeptide expre
42	1576.8	19.2	10367	15 AA058701	PIRB332. Syntheti
43	1568.4	19.1	2462	20 AAV82748	Pig endogenous ret
44	1550.4	18.9	7308	18 AAT90693	Plasmid hCMV+Intro
45	1550.4	18.9	7616	18 AAT90688	Plasmid Ceb coding

#### ALIGNMENTS

RESULT 1	AAV09700	standard; DNA: 8209 BP.
ID	AAV09700	standard; DNA: 8209 BP.
XX	AAV09700;	
XX	20-MAY-1998	(first entry)
DE	Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.	
XX	Porcine retrovirus: POEV; POL protein; ENV protein; GAG protein;	
KW	Virion core polypeptide; polymerase protein; envelope protein;	
KW	vacine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.	
XX	Porcine retrovirus.	
OS	Porcine retrovirus.	
FH	Key	Location/Qualifiers
FT	misc-feature	1..61
FT	/*tag- a	/note- "R LTR domain"
FT	1..588	/tag- b
FT	/*tag- b	62..143
FT	misc-feature	/*tag- c
FT	/*tag- c	/note- "U5 LTR domain"
FT	588..2162	/*tag- d
FT	/*tag- d	/product= GAG protein
FT	/*tag- e	/note- "viral core polypeptide"
FT	2163..5747	/*tag- e
FT	/product= POL protein	/note- "polymerase peptide"
FT	5620..7590	
FT	CDS	

```

FT FT /tag= f
FT FT /product= ENV protein
FT FT /note= "envelope protein"
FT FT 7591..8209
FT FT /tag= g
FT FT 7638..8106
FT FT /tag= h
FT FT /note= "u3 LTR domain"
FT FT 7643..7686
FT FT /tag= i
FT FT /note= "Binding site for transcription factors
FT FT OCT-1, c-Myb, LyF-1, E47 and ETS-1"
FT FT 7719..7739
FT FT /tag= j
FT FT /note= "Binding site for transcription factors
FT FT 7744..7754
FT FT /tag= k
FT FT /note= "Binding site for transcription factor AP-1/TR"
FT FT 7773..7785
FT FT /tag= l
FT FT /note= "Binding site for transcription factor ETS-1/GATA"
FT FT 7793..7834
FT FT /tag= m
FT FT /note= "Binding site for transcription factors
FT FT GATA, ETS-1, c-Myb, AP-1, GATA, E47"
FT FT 7841..7851
FT FT /tag= n
FT FT /note= "Binding site for transcription factor AP-1"
FT FT 7873..7883
FT FT /tag= o
FT FT /note= "Binding site for transcription factor AP-1"
FT FT 7885..7914
FT FT /tag= p
FT FT /note= "Binding site for transcription factors
FT FT AP-1/GATA, c-Myb and AP-1"
FT FT 7928..7953
FT FT /tag= q
FT FT /note= "Binding site for transcription factors E47
FT FT and ETS-1"
FT FT 7964..7968
FT FT /tag= r
FT FT 7985..7997
FT FT /tag= s
FT FT /note= "Binding site for transcription factor GATA"
FT FT 8003..8007
FT FT /tag= t
FT FT 8024..8036
FT FT /tag= u
FT FT /note= "Binding site for transcription factor GATA"
FT FT 8042..8046
FT FT /tag= v
FT FT 8049..8060
FT FT /tag= w
FT FT /note= "Binding site for transcription factor
FT FT AP-1/CREB"
FT FT 8075..8090
FT FT /tag= x
FT FT 1..61
FT FT /tag= y
FT FT /note= "R LTR domain"
FT FT 8127..8135
FT FT /tag= z
FT FT /note= "PolyA downstream element"
FT FT 8161..8166
FT FT /tag= aa
FT FT
PN WO9740167-A1.
XX
XX 30-OCT-1997.
PD
XX 18-APR-1997; 97WO-GB01087.
XX

```

```

PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMUT-) IMMURAN LTD.
PA (OONE-) O-ONE BIOTECH LTD.
PI Galbraith DN, Haworth C, Lees GW, Smith RT;
XX WPI; 1997-535851/49.
DR P-P-SDB; AAM39271; AAM39272; AAM39273.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT xeno-transplantation
XX
PS Claim 4; Fig 3; 69pp; English.
XX
CC This DNA sequence encodes the porcine retrovirus (POEV) virion core
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and
CC also includes the long terminal repeat (LTR). These proteins can be used
CC to develop viral vaccines, antisense nucleic acids, ribozymes and other
CC antiviral agents. They can also be used in xeno-transplantation
CC technology and as diagnostic tools.
XX
SQ Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

Query Match 100.0%; Score 8209; DB 18; Length 8209;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtagtgtaagctgtggcccgccgcttggaataaatacctctgtgttcac 60
DB 1 gtagtgtaagctgtggcccgccgcttggaataaatacctctgtgttcac 60
QY 61 aagaccgtcttcgtagtgatgttggtgtgtccttcgcagcccgagagggat 120
DB 61 aagaccgtcttcgtagtgatgttggtgtgtccttcgcagcccgagagggat 120
QY 121 tgtctttacgtgcttcatttgctgctgtggcggaatccttgacaccccta 180
DB 121 tgtctttacgtgcttcatttgctgctgtggcggaatccttgacaccccta 180
QY 181 caaccgagaaccgacttggaagtaaggatcccttggaagctgtgtgtcggcc 240
DB 181 caaccgagaaccgacttggaagtaaggatcccttggaagctgtgtgtcggcc 240
QY 241 ggcgtctctgttctagtgctgttttcgtagagcgcgcttcggttcagcttcc 300
DB 241 ggcgtctctgttctagtgctgttttcgtagagcgcgcttcggttcagcttcc 300
QY 301 ctcaagccgtaagagacttggaagctgtatcaagacagctgtctggaagacagctg 360
DB 301 ctcaagccgtaagagacttggaagctgtatcaagacagctgtctggaagacagctg 360
QY 361 ccaccttgaggagccgcccgagggtgggagagccagagcgctgtgttcctact 420
DB 361 ccaccttgaggagccgcccgagggtgggagagccagagcgctgtgttcctact 420
QY 421 gtccgtcagagagcagatctgtgttgaagcgaaagctccctccgcgcgcgcg 480
DB 421 gtccgtcagagagcagatctgtgttgaagcgaaagctccctccgcgcgcgcg 480
QY 481 acctttgctctgttggaagacgagcgaggtgcgtgtgtctgagactgtgttt 540
DB 481 acctttgctctgttggaagacgagcgaggtgcgtgtgtctgagactgtgttt 540
QY 541 ctgtctcgtgtgtctgtctgtcgttcctgtctacagtttaataatggagacag 600
DB 541 ctgtctcgtgtgtctgtctgtcgttcctgtctacagtttaataatggagacag 600
QY 601 tgactaccccttaagtttgaactctgcacattggagactgaattagatccaggctcata 660
DB 601 tgactaccccttaagtttgaactctgcacattggagactgaattagatccaggctcata 660

```



[illegible]

QY	1741	ggcagtcgcttcygtatatacagaagaactctcaagagctggaaggtgtacagagctg	1800
Db	1741	ggcagtcgcttcygtatatacagaagaactctcaagagctggaaggtgtacagagagctg	1800
QY	1801	agttacgagatctagtgagagagcgacagagaggtgtgtattacagaaagggagacagagag	1866
Db	1801	agttacgagatctagtgagagagcgacagagaggtgtgtattacagaaagggagacagagag	1866
QY	1861	agaaggaacagagaaagaaagagagagagaaagagagagagagagagagagagagag	1920
Db	1861	agaaggaacagagaaagaaagagagagagagagagagagagagagagagagagagag	1920
QY	1921	aagagaagaatttactaagatcttgccgcaatggtgttaagagggaagcagcagagaga	1986
Db	1921	aagagaagaatttactaagatcttgccgcaatggtgttaagagggaagcagcagagaga	1986
QY	1981	gagagagagagatttagaataaattagttcagagccctcagacaagtcagaggaacctgggcaata	2040
Db	1981	gagagagagagatttagaataaattagttcagagccctcagacaagtcagaggaacctgggcaata	2040
QY	2041	ggaaccccatctgacaagaagaccagttgctgattgttaagaaagaaagacacatctgacaaga	2100
Db	2041	ggaaccccatctgacaagaagaccagttgctgattgttaagaaagaaagacacatctgacaaga	2100
QY	2101	actgccccaaagagagaaacaaagagagagagagagagagagagagagagagagagagag	2160
Db	2101	actgccccaaagagagaaacaaagagagagagagagagagagagagagagagagagagag	2160
QY	2161	aggggagagagaggtctcgagacccctcccgagcccaaggttaacttgaaggttgagagggc	2220
Db	2161	aggggagagagaggtctcgagacccctcccgagcccaaggttaacttgaaggttgagagggc	2220
QY	2221	aaccagttgagttcttccttggttgaataccggagcggagagatctcaatgctgctacaaacattag	2280
Db	2221	aaccagttgagttcttccttggttgaataccggagcggagagatctcaatgctgctacaaacattag	2280
QY	2281	gaaacactaaagaaanaaaaaatctctgggtgagtggtgcacaaagggacacgagctatccat	2340
Db	2281	gaaacactaaagaaanaaaaaatctctgggtgagtggtgcacaaagggacacgagctatccat	2340
QY	2341	ggactacccggaagaagaccggttgaacttgagaggtgagacggglaaaccaactcgtttctgtca	2400
Db	2341	ggactacccggaagaagaccggttgaacttgagaggtgagacggglaaaccaactcgtttctgtca	2400
QY	2401	tcacctgagttgcgcagtaacccctctcagtgtagaacttactgcccagaagtgagagctcctaa	2460
Db	2401	tcacctgagttgcgcagtaacccctctcagtgtagaacttactgcccagaagtgagagctcctaa	2460
QY	2461	tttctcttggaacaaagagacccgaaggtgctctgataatacaaacccatcaactggttga	2520
Db	2461	tttctcttggaacaaagagacccgaaggtgctctgataatacaaacccatcaactggttga	2520
QY	2521	ccctccaattagatgatatacgaatctctcccaagcttcgccaagtlaagccttgatcaagaata	2580
Db	2521	ccctccaattagatgatatacgaatctctcccaagcttcgccaagtlaagccttgatcaagaata	2580
QY	2581	tacagtcctggttggagacgaatttcgcccaagcctgggacagaacccgacagagatgggttgg	2640
Db	2581	tacagtcctggttggagacgaatttcgcccaagcctgggacagaacccgacagagatgggttgg	2640
QY	2641	caaaagcaagttccccccacagttatccaactgaagggcaggtgtctacaccagttatcagfca	2700
Db	2641	caaaagcaagttccccccacagttatccaactgaagggcaggtgtctacaccagttatcagfca	2700
QY	2701	gacagttaccccttgagtagagaggtctcgagaaaggaatttggccgcagattgtcaaaagattaa	2760
Db	2701	gacagttaccccttgagtagagaggtctcgagaaaggaatttggccgcagattgtcaaaagattaa	2760
QY	2761	tcacaagggagatctcaggttcctgcgtccaatcccccttgagataatccccctgctacacgtta	2820
Db	2761	tcacaagggagatctcaggttcctgcgtccaatcccccttgagataatccccctgctacacgtta	2820

QY 2821 ggaagcctgggacaaatgatataatgacagctacagagcttgagagaggtcaataaagg 2880  
DB 2821 ggaagcctgggacaaatgatataatgacagctacagagcttgagagaggtcaataaagg 2880  
QY 2881 tgcagagacatacaaccaacaggtcccgaaaccttaacctcttaagccctccgcgctg 2940  
DB 2881 tgcagagacatacaaccaacaggtcccgaaaccttaacctcttaagccctccgcgctg 2940  
QY 2941 aacggaactgylacaacagatltggaactaaagaatgcctctctctgctgagatlaacc 3000  
DB 2941 aacggaactgylacaacagatltggaactaaagaatgcctctctctgctgagatlaacc 3000  
QY 3001 ccactagcccaacctcttctgcttgcaatgagagatccagylacagggaaagacccggc 3060  
DB 3001 ccactagcccaacctcttctgcttgcaatgagagatccagylacagggaaagacccggc 3060  
QY 3061 agctaacctgagccagactgcccgaaggttcaagaactcccgcgaacctcttaagaa 3120  
DB 3061 agctaacctgagccagactgcccgaaggttcaagaactcccgcgaacctcttaagaa 3120  
QY 3121 ccctacacagggagccttgccaaactcagatccaaacccctcaagylgacctctccagt 3180  
DB 3121 ccctacacagggagccttgccaaactcagatccaaacccctcaagylgacctctccagt 3180  
QY 3181 acgtgagatgacctctctgctggcgagcccaacaaacagagctgcttgagagatcagaag 3240  
DB 3181 acgtgagatgacctctctgctggcgagcccaacaaacagagctgcttgagagatcagaag 3240  
QY 3241 cactactgctggaattgltctgacctagctacagagcctctgctcaagaagccagattc 3300  
DB 3241 cactactgctggaattgltctgacctagctacagagcctctgctcaagaagccagattc 3300  
QY 3301 gcaagagagagatgaataactctgggtgacaglttgcgggcgggcagcgatgctgacg 3360  
DB 3301 gcaagagagagatgaataactctgggtgacaglttgcgggcgggcagcgatgctgacg 3360  
QY 3361 aggcacggaaagaaactgtagtccagatccagctcccaacaaacagcccaacaaatgtag 3420  
DB 3361 aggcacggaaagaaactgtagtccagatccagctcccaacaaacagcccaacaaatgtag 3420  
QY 3421 agtttttgggacagctgagatlttcgagactgtagtcccggttctgcaaccttaacg 3480  
DB 3421 agtttttgggacagctgagatlttcgagactgtagtcccggttctgcaaccttaacg 3480  
QY 3481 ccccaactctacccgttaaccaaaagaaaggggagttctctggtgctctgagcaacaga 3540  
DB 3481 ccccaactctacccgttaaccaaaagaaaggggagttctctggtgctctgagcaacaga 3540  
QY 3541 aggcatttgatgctatacaaaaggcctctgctgagcgacactgctctggtccctccgagc 3600  
DB 3541 aggcatttgatgctatacaaaaggcctctgctgagcgacactgctctggtccctccgagc 3600  
QY 3601 taactaaaccttacccttattgtgagatagcgttaagggagtagcccgagaggtttta 3660  
DB 3601 taactaaaccttacccttattgtgagatagcgttaagggagtagcccgagaggtttta 3660  
QY 3661 cccaacccctagagcaatgagagagactgctgctacactgcaaaagaactgtgacctg 3720  
DB 3661 cccaacccctagagcaatgagagagactgctgctacactgcaaaagaactgtgacctg 3720  
QY 3721 taagccagttgtggcccgatgtctgaagctgtgcaagctgtgagcacaatgtgcaagg 3780  
DB 3721 taagccagttgtggcccgatgtctgaagctgtgcaagctgtgagcacaatgtgcaagg 3780  
QY 3781 acgctgcaaatgtgactttgagacagataatactgtaatagcccccatgtgagag 3840  
DB 3781 acgctgcaaatgtgactttgagacagataatactgtaatagcccccatgtgagag 3840  
QY 3841 acatgctgagcgccccagacagatgagatgacaaagcccgatgacccaatcaataa 3900  
DB 3841 acatgctgagcgccccagacagatgagatgacaaagcccgatgacccaatcaataa 3900  
QY 3901 gctgctctccacagagaggtcaacttctgctccacacgctctcaacctgccaatc 3960

DB 3901 gctgctctccacagagaggtcaacttctgctccacacgctctcaacctgccaatc 3960  
QY 3961 tctgctctgagagactgtagaacagtagtactaataatgtgccaatctgtagag 4020  
DB 3961 tctgctctgagagactgtagaacagtagtactaataatgtgccaatctgtagag 4020  
QY 4021 agactgggtgcgcaagacacttacagataccgctgactgagagagtgtaacctggt 4080  
DB 4021 agactgggtgcgcaagacacttacagataccgctgactgagagagtgtaacctggt 4080  
QY 4081 tcaactagcggaaacagctatgtgtgagaggttaagagatgctgggctggcgagtggt 4140  
DB 4081 tcaactagcggaaacagctatgtgtgagaggttaagagatgctgggctggcgagtggt 4140  
QY 4141 acgggaccccgacagatctggccagagcctccggaagaagacttcaagcgcaaaagctg 4200  
DB 4141 acgggaccccgacagatctggccagagcctccggaagaagacttcaagcgcaaaagctg 4200  
QY 4201 agctatggcctcagcgcaagcttggcgctggccgaaggaatccataaacttata 4260  
DB 4201 agctatggcctcagcgcaagcttggcgctggccgaaggaatccataaacttata 4260  
QY 4261 cggagacaggtatgtccttggcactgacacagtaacggggccatctataaacaagg 4320  
DB 4261 cggagacaggtatgtccttggcactgacacagtaacggggccatctataaacaagg 4320  
QY 4321 ggttcttaccctcagagagaggaataaagaacaaagaagaaattcaagccattag 4380  
DB 4321 ggttcttaccctcagagagaggaataaagaacaaagaagaaattcaagccattag 4380  
QY 4381 aagccttaacttgcacaaagagctagctatatacaactgtccctgagacatcagaagcca 4440  
DB 4381 aagccttaacttgcacaaagagctagctatatacaactgtccctgagacatcagaagcca 4440  
QY 4441 aagatctatctagaaggaacccagatgctgcagccgggttgcgaagcagagcccg 4500  
DB 4441 aagatctatctagaaggaacccagatgctgcagccgggttgcgaagcagagcccg 4500  
QY 4501 ctgttaaccttgccttaataaagaacgccccaaagccccaaacccagacagatca 4560  
DB 4501 ctgttaaccttgccttaataaagaacgccccaaagccccaaacccagacagatca 4560  
QY 4561 cccctagaagactgcgcaagagataaaagaatagacaggtctctgagactccggagggga 4620  
DB 4561 cccctagaagactgcgcaagagataaaagaatagacaggtctctgagactccggagggga 4620  
QY 4621 cctgctataccctcataltggaaggaatctcctcccaaaagaaggttagaatagtcc 4680  
DB 4621 cctgctataccctcataltggaaggaatctcctcccaaaagaaggttagaatagtcc 4680  
QY 4681 aacagatcatgcttaaccccaactaggaactaaacacactgagagcttgcagaat 4740  
DB 4681 aacagatcatgcttaaccccaactaggaactaaacacactgagagcttgcagaat 4740  
QY 4741 cccctatcatgttctgaggtacacagaggtgctgacatcggtgtgtcaaaactgtgctc 4800  
DB 4741 cccctatcatgttctgaggtacacagaggtgctgacatcggtgtgtcaaaactgtgctc 4800  
QY 4801 cctgccaagctgttgaatgttaactctccagaataccctccaggaagaagactaagggaa 4860  
DB 4801 cctgccaagctgttgaatgttaactctccagaataccctccaggaagaagactaagggaa 4860  
QY 4861 gccacccaagcgctacatgagaggtggaacttcaactgagagtaaaagccgcttaaacgaa 4920  
DB 4861 gccacccaagcgctacatgagaggtggaacttcaactgagagtaaaagccgcttaaacgaa 4920  
QY 4921 acaaatatcatgtgttctgtagacacacttctcaggtatggatagaggtctataacta 4980  
DB 4921 acaaatatcatgtgttctgtagacacacttctcaggtatggatagaggtctataacta 4980  
QY 4981 agaaagagacttcaacggtgtgtgtaagaaatactgagagaaatttcccaagattg 5040

Db 4981 agaaagagacttcaacgttggctaaagaalacttggagaaattttlccaagatttg 5040  
QY 5041 gaatacctaagtaataaggtcagaacaatgtgccagcttctgttgcacagtaagtacag 5100  
Db 5041 gaatacctaagtaataaggtcagaacaatgtgccagcttctgttgcacagtaagtacag 5100  
QY 5101 gacttgcacaagatatltgggagattgtatgtaaacctgtcatltgtcaataagcccaagct 5160  
Db 5101 gacttgcacaagatatltgggagattgtatgtaaacctgtcatltgtcaataagcccaagct 5160  
QY 5161 caggaacagtagtagagagatgtatgaacacataaagaagacccttaaccaaatltgacacag 5220  
Db 5161 caggaacagtagtagagagatgtatgaacacataaagaagacccttaaccaaatltgacacag 5220  
QY 5221 agactggcattaaatgattgtagtctccctgccttctgtcctttttagaggtagagaca 5280  
Db 5221 agactggcattaaatgattgtagtctccctgccttctgtcctttttagaggtagagaca 5280  
QY 5281 ccccttgcagacttctgggctgcaccccttaataatgtcatcaggggagaccccccgttgg 5340  
Db 5281 ccccttgcagacttctgggctgcaccccttaataatgtcatcaggggagaccccccgttgg 5340  
QY 5341 caggaatattgcctttgcacacatagctgtatgtctgtcttccacagcttcttctcagtc 5400  
Db 5341 caggaatattgcctttgcacacatagctgtatgtctgtcttccacagcttcttctcagtc 5400  
QY 5401 tcaagagcgtctcgaatgggtgtagagcagcagctgtgaagcagcttccggagagccctacag 5460  
Db 5401 tcaagagcgtctcgaatgggtgtagagcagcagctgtgaagcagcttccggagagccctacag 5460  
QY 5461 gaggaagacttgcagacttccacatctgccttccaagtttgagatltcaatgtatlagacgc 5520  
Db 5461 gaggaagacttgcagacttccacatctgccttccaagtttgagatltcaatgtatlagacgc 5520  
QY 5521 accgttgaggaagaaccttcgaactcgtgtggaaggagaccttcttcgtacttttgaacacac 5580  
Db 5521 accgttgaggaagaaccttcgaactcgtgtggaaggagaccttcttcgtacttttgaacacac 5580  
QY 5581 caacggtctgtgaagaagtcgaagaaatccccccttgatccatctgcatcccaagttlaagccgg 5640  
Db 5581 caacggtctgtgaagaagtcgaagaaatccccccttgatccatctgcatcccaagttlaagccgg 5640  
QY 5641 cgcacactccgcagctcgggtgtgtagagcgaagaagacttgaatctcccttaagcttcgc 5700  
Db 5641 cgcacactccgcagctcgggtgtgtagagcgaagaagacttgaatctcccttaagcttcgc 5700  
QY 5701 tccatcgccttggttcccttaacttaacaaataactcccccagccagtagtaaacgcttata 5760  
Db 5701 tccatcgccttggttcccttaacttaacaaataactcccccagccagtagtaaacgcttata 5760  
QY 5761 gacagctcgaacccccaataacacttatacccttaactcctgattatltgacccctgatacg 5820  
Db 5761 gacagctcgaacccccaataacacttatacccttaactcctgattatltgacccctgatacg 5820  
QY 5821 ggtgtcaacttgaatagacactcgaagtgttgcctcctagagacactgtgtgacctgaactg 5880  
Db 5821 ggtgtcaacttgaatagacactcgaagtgttgcctcctagagacactgtgtgacctgaactg 5880  
QY 5881 catttcgcctccgattgataaccccgctgttaaaaagacacccctcccaactcagtcgct 5940  
Db 5881 catttcgcctccgattgataaccccgctgttaaaaagacacccctcccaactcagtcgct 5940  
QY 5941 agtataaggttctatctgtcgcacagacagagagaagaataactgttggggttcttgag 6000  
Db 5941 agtataaggttctatctgtcgcacagacagagagaagaataactgttggggttcttgag 6000  
QY 6001 gaatccttctgttagagatgtgaactcgttcaactccaagatgtgaacttggaaatgtccgg 6060  
Db 6001 gaatccttctgttagagatgtgaactcgttcaactccaagatgtgaacttggaaatgtccgg 6060  
QY 6061 atctcttcacagagccgggttaaaatctctctgttcaatccggccgggcaagtacaaa 6120  
Db 6061 atctcttcacagagccgggttaaaatctctctgttcaatccggccgggcaagtacaaa 6120

QY 6121 atgtataactatataagaataagactgtctccccaacagacttaattatctaaagata 6180  
Db 6121 atgtataactatataagaataagactgtctccccaacagacttaattatctaaagata 6180  
QY 6181 agttcaactgaagaagaagaataataattcaaaagtgtatgaatgtatgtgcgtggg 6240  
Db 6181 agttcaactgaagaagaagaataataattcaaaagtgtatgaatgtatgtgcgtggg 6240  
QY 6241 aatagtttttatataatgtcggggagcagaggttccaattlaaccaatcgccttaagata 6300  
Db 6241 aatagtttttatataatgtcggggagcagaggttccaattlaaccaatcgccttaagata 6300  
QY 6301 gagacggggagacaaacccctgtggaatgtggaccccgataaagtactgtgtgaacgggg 6360  
Db 6301 gagacggggagacaaacccctgtggaatgtggaccccgataaagtactgtgtgaacgggg 6360  
QY 6361 ccccgcccttgcagcagcccgataacttgcggttgcaccccaatlaacctgcgtgcgct 6420  
Db 6361 ccccgcccttgcagcagcccgataacttgcggttgcaccccaatlaacctgcgtgcgct 6420  
QY 6421 gacataacacagcgccttagcaacagttaccactgtatltccttaaccaacagcctaga 6480  
Db 6421 gacataacacagcgccttagcaacagttaccactgtatltccttaaccaacagcctaga 6480  
QY 6481 aactcccgaggtgttctctgttgaagacagagagacttcaagtctcaacaggagct 6540  
Db 6481 aactcccgaggtgttctctgttgaagacagagagacttcaagtctcaacaggagct 6540  
QY 6541 ttccaagcatcaactccaccccgacttgaatgtccacttcttctgtgtgtctgtctatcc 6600  
Db 6541 ttccaagcatcaactccaccccgacttgaatgtccacttcttctgtgtgtctgtctatcc 6600  
QY 6601 tcaaggcccttataatgaagggtatgtcttaagaagaataatltcaatgtgaccaaagag 6660  
Db 6601 tcaaggcccttataatgaagggtatgtcttaagaagaataatltcaatgtgaccaaagag 6660  
QY 6661 catagaataactgtatactgggggttccggaataagcttaacctcaactgaatltccggg 6720  
Db 6661 catagaataactgtatactgggggttccggaataagcttaacctcaactgaatltccggg 6720  
QY 6721 aaggggacatgtataggaagaagctcccccataccccaacacacttgcctatagctgtg 6780  
Db 6721 aaggggacatgtataggaagaagctcccccataccccaacacacttgcctatagctgtg 6780  
QY 6781 gtttaagacaggtcctcaaaaaatacagttatagtaactgtgtataacaggtgtgtggca 6840  
Db 6781 gtttaagacaggtcctcaaaaaatacagttatagtaactgtgtataacaggtgtgtggca 6840  
QY 6841 tgaataactgggttlaaccccgctgttccacactcagttcaacaaatccaagaatttc 6900  
Db 6841 tgaataactgggttlaaccccgctgttccacactcagttcaacaaatccaagaatttc 6900  
QY 6901 tgtgtcatgttccaaatcgttcccccaggtgttactaacactcctgtgagaggttgccttgat 6960  
Db 6901 tgtgtcatgttccaaatcgttcccccaggtgttactaacactcctgtgagaggttgccttgat 6960  
QY 6961 gaatatgactatcgttataacacgacacaaaagaagaccogtatacccttaacctgaactga 7020  
Db 6961 gaatatgactatcgttataacacgacacaaaagaagaccogtatacccttaacctgaactga 7020  
QY 7021 atgcctcgaattagagagacgctgttgcgttaagaacaaggagacgctgccttgatcaaga 7080  
Db 7021 atgcctcgaattagagagacgctgttgcgttaagaacaaggagacgctgccttgatcaaga 7080  
QY 7081 ccaacagcagcttagagaaagagacttggtagactcaatctgcgcacatgtacagaagatctccga 7140  
Db 7081 ccaacagcagcttagagaaagagacttggtagactcaatctgcgcacatgtacagaagatctccga 7140  
QY 7141 gctttaaagaggtctgttagcaacctagaagaagttcccgacttcttctggaaggtgt 7200  
Db 7141 gctttaaagaggtctgttagcaacctagaagaagttcccgacttcttctggaaggtgt 7200

OY	7201	ctacaggaacccggaaggattagatgccttcgctttcttaagaagagtggtttatgtgcaacc	7260
Db	7201	ctacaggaacccggaaggattatagatcctgcctgtttctaaaggaagtggtttatgtgcacc	7260
OY	7261	ttaaagaagaatgtgtctctctatgtatagatcgaacgaagccatcagaagactcatgaac	7320
Db	7261	ttaaagaagaagaatgtgtctctctatgtatagatcgaacgaagccatcagaagactcatgaac	7320
OY	7321	aagcttgaaaaaagatttagaggaagcgttcgaagggaaggaagcctgacacagggtgtgtt	7380
Db	7321	aagcttcgaaaaaagatttagaggaagcgttcgaagggaaggaagcctgacacagggtgtgtt	7380
OY	7381	gaagatcgtgttcaacagctctccttgatgataccaccctgtcttctgcctcgaagggcc	7440
Db	7381	gaagatcgtgttcaacagctctccttgatgataccaccctgtcttctgcctcgaagggcc	7440
OY	7441	ctatagatcctgcctcctgttacttaacaagttgggccttcgtcttaataatagttgtttgc	7500
Db	7441	ctatagatcctgcctcctgttacttaacaagttgggccttcgtcttaataatagttgtttgc	7500
OY	7501	tttgttgaagacagagtgagtgagtcgacccgaatcatgtatgtacttaggcacaaatlaacgaagc	7560
Db	7501	tttgttgaagacagagtgagtgagtcgacccgaatcatgtatgtacttaggcacaaatlaacgaagc	7560
OY	7561	ctctcgtgccaagaggaagaacttgacccctcgaagcttccacgttctaagatlaagaactataa	7620
Db	7561	ctctcgtgccaagaggaagaacttgacccctcgaagcttccacgttctaagatlaagaactataa	7620
OY	7621	caagaacagaagtggtggaatggaagaagtgaaaaaatgaaacttaacccctccagaacccag	7680
Db	7621	caagaacagaagtggtggaatggaagaagtgaaaaaatgaaacttaacccctccagaacccag	7680
OY	7681	aagttataaaaaagcctctaataatgcccccggaattacagaccctgcgtcgtccgaataata	7740
Db	7681	aagttataaaaaagcctctaataatgcccccggaattacagaccctgcgtcgtccgaataata	7740
OY	7741	ggttgaagaagttcaacacttccattgtctccaggcctgctatccctggcctaagtgaataaac	7800
Db	7741	ggttgaagaagttcaacacttccattgtctccaggcctgctatccctggcctaagtgaataaac	7800
OY	7801	aggaagaatgagttgactaatcgtctatcctgatatctgtatctgtataaactgtgacccataagaag	7860
Db	7801	aggaagaatgagttgactaatcgtctatcctgatatctgtatctgtataaactgtgacccataagaag	7860
OY	7861	aattgatthacaacatlgacagccctagtgacactatcctcaactcgaactcgtcaactctgcgc	7920
Db	7861	aattgatthacaacatlgacagccctagtgacactatcctcaactcgaactcgtcaactctgcgc	7920
OY	7921	aggagcccaacgacatgcccgaacctccggagcctatttaaatgattgtgtccaggaagc	7980
Db	7921	aggagcccaacgacatgcccgaacctccggagcctatttaaatgattgtgtccaggaagc	7980
OY	7981	gggctctcgatattttaaatagtatgtgtccatcgaagcgggctctcgatattttaaat	8040
Db	7981	gggctctcgatattttaaatagtatgtgtccatcgaagcgggctctcgatattttaaat	8040
OY	8041	gattggttttgtagagccagcgtttgttbtgaaaccccataaagaactgtgccgattccgca	8100
Db	8041	gattggttttgtagagccagcgtttgttbtgaaaccccataaagaactgtgccgattccgca	8100
OY	8101	ctcgggcccgaagccctcctacccctcggctggtgtgaagactgtgggcccacgacgctttgg	8160
Db	8101	ctcgggcccgaagccctcctacccctcggctggtgtgaagactgtgggcccacgacgctttgg	8160
OY	8161	aataaaaatccctctgctgtttgcatcaaaaaaanaaaaaaanaa 8209	
Db	8161	aataaaaatccctctgctgtttgcatcaaaaaaanaaaaaaanaa 8209	
RESULT 2			
AAV09699 standard; DNA: 8196 BP.			
ID AAV09699			
XX			

AC	AAV09699;
XX	
DT	19-MAY-1998 (first entry)
XX	
DE	Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX	
KM	Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;
KM	vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX	
OS	Porcine retrovirus.
XX	
Key	Location/Qualifiers
FT	576..2126
FT	/*tag= a
FT	/product= GAG protein
FT	/note= "virlion core polypeptide"
FT	2143..5733
FT	/*tag= b
FT	/product= POL protein
FT	/note= "polymerase peptide sequence as given in specification"
FT	5606..7576
FT	/*tag= c
FT	/product= ENV protein
FT	/note= "envelope protein"
PN	W09740167-A1.
XX	
PD	30-OCT-1997.
XX	
PF	18-APR-1997; 97WO-GB01087.
XX	
PR	10-FEB-1997; 97GB-0002668.
PR	19-APR-1996; 96GB-0008164.
XX	
PA	(IMOT-) IMOTRAN LTD.
PA	(OONE-) O-ONE BIOTECH LTD.
XX	
PI	Galbraith DN, Haworth C, Lees GM, Smith KT;
XX	
DR	WPI, 1997-535851/49.
XX	
PT	Polynucleotide encoding porcine retrovirus expression product -
PT	useful to develop products for use in vaccines, diagnosis and
PT	xeno-transplantation
PS	Claim 4; Fig 2; 69pp; English.
XX	
CC	This DNA sequence encodes the porcine retrovirus (PoEV) virlion core
CC	polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC	proteins can be used to develop viral vaccines, antisense nucleic acids,
CC	ribozymes and other antiviral agents. They can also be used in
CC	xeno-transplantation technology and as diagnostic tools.
XX	
Sequence	8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other:
Query Match	99.5%; Score 8171.8; DB 18; Length 8196;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 8192; Conservative	3; Mismatches 0; Indels 2; Gaps 2
13	tgtggagcccaagcgcgttggaataaatacctctctgtctgtttgcatcaagacgcttct 72
1	tgtggagcccaagcgcgttggaataaatacctctctgtctgtttgcatcaagacgcttct 60
73	cgtgagtgattgggggtgtcgcctctccagaccgcgagcaggggattgtcttcttact 132
61	cgtgagtgattgggggtgtcgcctctccagaccgcgagcaggggattgtcttcttact 120
133	ggcccttaattggtgctgtggccgggaaatctctcgaccaccccttaacccgaaaac 192
121	ggcccttaattggtgctgtggccgggaaatctctcgaccaccccttaacccgaaaac 180

QY	193	gacttgcagagctaaagagatccccccttttgaaacgtgtgtgtgtgtcgtccgagcgctctcgtt	252
Db	181	gacttgcagagctaaagagatccccccttttgaaacgtgtgtgtgtgtcgtccgagcgctctcgtt	240
QY	253	ctgaaatgtctcttttcggatctgagagcgcttcgttgaaagctgtctctctcaagacgttaa	312
Db	241	ctgaaatgtctcttttcggatctgagagcgcttcgttgaaagctgtctctctcaagacgttaa	300
QY	313	ggacactgagagacatctgatacaacaacgtctgataagaaatcacagcgctgtccacccctggggg	372
Db	301	ggacactgagagacatctgatacaacaacgtctgataagaaatcacagcgctgtccacccctggggg	360
QY	373	acgcccgcgagcgctgagggagagccagggagccgctgtgtgtctctactctgtctgtctcaagag	432
Db	361	acgcccgcgagcgctgagggagagccagggagccgctgtgtgtgtctctactctgtctcaagag	420
QY	433	accgagatctcttttttaagaggaaagttccccctccgcgcgcgcgcgcgcgcgcgcgcgcgcgc	492
Db	421	accgagatctcttttttaagaggaaagttccccctccgcgcgcgcgcgcgcgcgcgcgcgcgcgc	480
QY	493	gcttctgagaaagacgagacgagctgcgcgtgtctgtgatccgtgtgtctgtctgtctgtctgtgtg	552
Db	481	gcttctgagaaagacgagacgagctgcgcgtgtctgtgatccgtgtgtctgtgtctgtctgtgtg	540
QY	553	tcttctgtctgt	612
Db	541	tcttctgtctgt	600
QY	613	ttaatttggaaacttcgaaactttgaaacttggaaacttggaaacttcgaaacttcgaaacttcgaaact	672
Db	601	ttaatttggaaacttcgaaactttgaaacttggaaacttggaaacttcgaaacttcgaaacttcgaaact	660
QY	673	aggttaagaagagacacttctgtgcgaacttctgtgcctctcctaalgccaaatccagatctgtgtg	732
Db	661	aggttaagaagagacacttctgtgcgaacttctgtgcctctcctaalgccaaatccagatctgtgtg	720
QY	733	gattgtccaatcaagagggagacctttaatctgtaaatatctctgtgcgtgttaagggcaatcaatt	792
Db	721	gattgtccaatcaagagggagacctttaatctgtaaatatctctgtgcgtgttaagggcaatcaatt	780
QY	793	ttcagaaactggagcccgagctcttcacatccgtatcagaagagcccatcttcctaagctgtgcaagatt	852
Db	781	ttcagaaactggagcccgagctcttcacatccgtatcagaagagcccatcttcctaagctgtgcaagatt	840
QY	853	tgtgcagaagatcctccgcgcatactgtgtttaaaacaaatgtgcataataaacaagaagaagcgagtc	912
Db	841	tgtgcagaagatcctccgcgcatactgtgtttaaaacaaatgtgcataataaacaagaagaagcgagtc	900
QY	913	cccgaaatccctgtgctcttctgtgagagaaaaaacaacaatccctgtgcgaaagaatctgagccctctt	972
Db	901	cccgaaatccctgtgctcttctgtgagagaaaaaacaacaatccctgtgcgaaagaatctgagccctctt	960
QY	973	cctcgtatctactaaccccgagatccagagggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1032
Db	961	cctcgtatctactaaccccgagatccagagggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1020
QY	1033	ccaaaccccttatccaaagacaagaggtgtgtgtgtgagggagacccctcgcgcgcgcgcgcgcgcgcgc	1092
Db	1021	ccaaaccccttatccaaagacaagaggtgtgtgtgtgagggagacccctcgcgcgcgcgcgcgcgcgcgc	1080
QY	1093	tgtgtgagagagacctgtctgtccgaggaactcggagccggagagagcgccaaaccccgagagcgagag	1152
Db	1081	tgtgtgagagagacctgtctgtccgaggaactcggagccggagagagcgccaaaccccgagagcgagag	1140
QY	1153	acgagatccgagatattaacgcctgtgcgaactatgtgcctccatgccagaggggccaattgtc	1212
Db	1141	acgagatccgagatattaacgcctgtgcgaactatgtgcctccatgccagaggggccaattgtc	1200
QY	1213	agccccctcagatattgcccccttctctctgcagaaatctcctaattgtggaataactaacctc	1272
Db	1201	agccccctcagatattgcccccttctctctgcagaaatctcctaattgtggaataactaacctc	1260
QY	1273	cccccttctcgcagagatccccaacgcgtccaaacgaggttgtgtgtgagctcccttatgttctctc	1332

[illegible]

Db 2339 gaacgcttgagctggagctggagcgggtaaacccacgctgttctggtlcatccctgagtc 2398  
QY 2413 cagtaaccctctagctgtagagactactgacccaagatgagagctccaattctttgaa 2472  
Db 2339 cagtaaccctctagctgtagagactactgacccaagatgagagctccaattctttgaa 2458  
QY 2473 aaggaagacccaagatgctgtgtgtaatacaaacccactactgtgtgtgacccccaatag 2532  
Db 2459 aaggaagacccaagatgctgtgtgtaatacaaacccactactgtgtgtgacccccaatag 2518  
QY 2533 atgataatgatacgtatattcttccccaagttaagccttgataagataatcagtcctgt 2592  
Db 2519 atgataatgatacgtatattcttccccaagttaagccttgataagataatcagtcctgt 2578  
QY 2593 tggagcagcttccccaagcctggcgagaaacccgaggaatgaggttttgccaagaatc 2652  
Db 2579 tggagcagcttccccaagcctggcgagaaacccgaggaatgaggttttgccaagaatc 2638  
QY 2653 ccccaaggtatatacactgaaagcagtgctacacacagatacagtlcagaacagtaaccct 2712  
Db 2639 ccccaaggtatatacactgaaagcagtgctacacacagatacagtlcagaacagtaaccct 2698  
QY 2713 tgaataagagagctcgaagaagaaattggcgcaatgttcaagaattatccacaaggca 2772  
Db 2699 tgaataagagagctcgaagaagaaattggcgcaatgttcaagaattatccacaaggca 2758  
QY 2773 tccatgttccctgccaatccctctggaatactccctctgatacaggttagaagaactgga 2832  
Db 2759 tccatgttccctgccaatccctctggaatactccctctgatacaggttagaagaactgga 2818  
QY 2833 ccaatgattatcgaacccagtaacagacttgagagaagtcataaagaagtgcaagacaac 2892  
Db 2819 ccaatgattatcgaacccagtaacagacttgagagaagtcataaagaagtgcaagacaac 2878  
QY 2893 acccaagcgtccgaacccctataactcttgagcgccctcccgccggaacggaactggt 2952  
Db 2879 acccaagcgtccgaacccctataactcttgagcgccctcccgccggaacggaactggt 2938  
QY 2953 acacagatattgagacttaaaagatgcttcttctgctgagattacacccactagcaac 3012  
Db 2939 acacagatattgagacttaaaagatgcttcttctgctgagattacacccactagcaac 2998  
QY 3013 cacttttgcttgaatactgagagatacaggtacggaagaacccgggcaactcagga 3072  
Db 2999 cacttttgcttgaatactgagagatacaggtacggaagaacccgggcaactcagga 3058  
QY 3073 cccgactgcccgaaggggtcaagaactcccgacacatcttgaacgaacccctacacaggg 3132  
Db 3059 cccgactgcccgaaggggtcaagaactcccgacacatcttgaacgaacccctacacaggg 3118  
QY 3133 accggtgcaactttagatccaacacccctcaggtgacccctcccgtaagtgatgagcc 3192  
Db 3119 accggtgcaactttagatccaacacccctcaggtgacccctcccgtaagtgatgagcc 3178  
QY 3193 tgcctctggcgagacccaacaacagagctgcttagaaggtaacgaagcactactgctg 3252  
Db 3179 tgcctctggcgagacccaacaacagagctgcttagaaggtaacgaagcactactgctg 3238  
QY 3253 aattgtctgacactagctacagagcctctgcttaagaagccacagatttgcaagagaggg 3312  
Db 3239 aattgtctgacactagctacagagcctctgcttaagaagccacagatttgcaagagaggg 3298  
QY 3313 taacatacttgagggtacaggtttggggcgagcgatgctgacggaagcagcgaga 3372  
Db 3299 taacatacttgagggtacaggtttggggcgagcgatgctgacggaagcagcgaga 3358  
QY 3373 aaactgtatgccaatacggcccaacacagccaacacagtgagagagtttttggga 3432  
Db 3359 aaactgtatgccaatacggcccaacacagccaacacagtgagagagtttttggga 3418  
QY 3433 cagctgagatttgacagactgtgataccggggtttgcgaccttagcagcccaactaac 3492  
Db 3419 cagctgagatttgacagactgtgataccggggtttgcgaccttagcagcccaactaac 3478

---

QY 3493 cgttaaccaaagaaagggggtatctccctgggctccttgagcaccagaagcatttgat 3552  
Db 3479 cgttaaccaaagaaagggggtatctccctgggctccttgagcaccagaagcatttgat 3538  
QY 3553 ctatacaaaagcctgtctgagcgacactgtctgtgcccctctgacgttaactaaacct 3612  
Db 3539 ctatacaaaagcctgtctgagcgacactgtctgtgcccctctgacgttaactaaacct 3598  
QY 3613 ttacccttatgtgagtgagcgtaaggagtgcccggaaggttttaacccaacactag 3672  
Db 3599 ttacccttatgtgagtgagcgtaaggagtgcccggaaggttttaacccaacactag 3658  
QY 3673 gacacatgagagagacctgttgcctacactgtcaaaagatltgactctgtaagcagtggt 3732  
Db 3659 gacacatgagagagacctgttgcctacactgtcaaaagatltgactctgtaagcagtggt 3718  
QY 3733 ggcctgtatgtctgaaagctatcgacagctgtgtgcaatactgttcaagagcctgcaaat 3792  
Db 3719 ggcctgtatgtctgaaagctatcgacagctgtgtgcaatactgttcaagagcctgcaaat 3778  
QY 3793 tgaacttgagacagaataactgttaataagccccaatgcatltgagacaatcgttcggc 3852  
Db 3779 tgaacttgagacagaataactgttaataagccccaatgcatltgagacaatcgttcggc 3838  
QY 3853 agccccaagcagatgagatgacccaacgcccgaatgacccaatacaaaagcctgtctca 3912  
Db 3839 agccccaagcagatgagatgacccaacgcccgaatgacccaatacaaaagcctgtctca 3898  
QY 3913 cagaagaggtactacttgcctccacagcgccttaaacctgtgcaacttctgctcctgag 3972  
Db 3899 cagaagaggtactacttgcctccacagcgccttaaacctgtgcaacttctgctcctgag 3958  
QY 3973 agactatgaacacagtgactatgatactgacataactltgataagagagactgggccc 4032  
Db 3959 agactatgaacacagtgactatgatactgacataactltgataagagagactgggccc 4018  
QY 4033 gcaagagccttaacagataccgctgacttgagaagtgctaaactgttcaactgacgaa 4092  
Db 4019 gcaagagccttaacagataccgctgacttgagaagtgctaaactgttcaactgacgaa 4078  
QY 4093 gcaagatgtgtgtggaagtgagaagatgctgagcgagcaatgtgagaggaacccga 4152  
Db 4079 gcaagatgtgtgtggaagtgagaagatgctgagcgagcaatgtgagaggaacccga 4138  
QY 4153 cgactgagccagcagcctgtccggaaggaacttcagcgcaaaagcctgagctcagccc 4212  
Db 4139 cgactgagccagcagcctgtccggaaggaacttcagcgcaaaagcctgagctcagccc 4198  
QY 4213 tcaaggaacttgcgctgagcgcgaaggaatccataaacttatacgcgaacaggt 4272  
Db 4199 tcaaggaacttgcgctgagcgcgaaggaatccataaacttatacgcgaacaggt 4258  
QY 4273 atgacttgcgactgacacagtaacagcgagccatctataaacaaggggtgttcaact 4332  
Db 4259 atgacttgcgactgacacagtaacagcgagccatctataaacaaggggtgttcaact 4318  
QY 4333 cagcaggggaggaataaagaacaaaggaatcttaagccttataaagccttaact 4392  
Db 4319 cagcaggggaggaataaagaacaaaggaatcttaagccttataaagccttaact 4378  
QY 4393 tgcacaaagcctgactatatacactgtcccgagacatcagaagccaaagatccat 4452  
Db 4379 tgcacaaagcctgactatatacactgtcccgagacatcagaagccaaagatccat 4438  
QY 4453 ctgagaggaacagatgctgacgggtgtgcaacagcagcagccagcgtttaaacttc 4512  
Db 4439 ctgagaggaacagatgctgacgggtgtgcaacagcagcagccagcgtttaaacttc 4498  
QY 4513 tgcctataataaagcccaagcccaagcccaagcccaagcagacagacacccctaagagct 4572  
Db 4499 tgcctataataaagcccaagcccaagcccaagcccaagcagacagacacccctaagagct 4558





Dh	6719	atggagaaagctcccccacatccacaaacacaccttgctatagctacgtgtgtttatgacga	6778
Qy	6793	gctccagaagaatcagatatttaagtaacctggtttataacaggtgtgtgtgacatgcag	6852
Dh	6779	gctctagaagaatacagatatttaagtaacctggtttataacaggtgtgtgtgacatgcag	6838
Qy	6853	ttaaaccccttggtttcacctcaagcttcctaaacatccaaagatttctgtgtcatgttc	6912
Dh	6839	ttaaacccctgtgtttcacctcaagcttcctaaacatccaaagatttctgtgtcatgttc	6898
Qy	6913	caaatgctcccccgagtgataccatccctcgaagaaagtgttcctgtatgaataatacat	6972
Dh	6899	caaatgctcccccgagtgataccatccctcgaagaaagtgttcctgtatgaataatacat	6958
Qy	6973	cggtataacacgaaccaaaagaagaaccgtatcccttaacccctgaatgtctcggatla	7032
Dh	6959	cggtataacacgaaccaaaagaagaaccgtatcccttaacccctgaatgtctcggatla	7018
Qy	7033	ggagccggccgtttgctgctgagaaagaaggaagccttcgtatatacagagacacagacta	7092
Dh	7019	ggagccggccgtttgctgctgagaaagaaggaagccttcgtatatacagagacacagacta	7078
Qy	7093	gagaaagagacttggctgagctacatactgggcacatgacacgaagatctccggaccttaagag	7152
Dh	7079	gagaaagagacttggctgagctacatactgggcacatgacacgaagatctccggaccttaagag	7138
Qy	7153	tcgtttgcaacctagaagaagctccctgacttctgtctgaagtggtctctacagaaccgg	7212
Dh	7139	tcgtttgcaacctagaagaagctccctgacttctgtctgaagtggtctctacagaaccgg	7198
Qy	7213	aggagatataatcgcgcgtttcttaagaagaagcttggtttatgtgcagaccttaagaagaa	7272
Dh	7199	aggagatataatcgcgcgtttcttaagaagaagcttggtttatgtgcagaccttaagaagaa	7258
Qy	7273	tgctgtcctctatgtatagatccatccagagagccatcaagaatccatgaacaagcttagaaa	7332
Dh	7259	tgctgtcctctatgtatagatccatccagagagccatcaagaatccatgaacaagcttagaaa	7318
Qy	7333	aagttagaagaagcgctgcgaagaaggaagaagcgctgcagcagggtgtgttgaagaatgcttc	7392
Dh	7319	aagttagaagaagcgctgcgaagaaggaagaagcgctgcagcagggtgtgttgaagaatgcttc	7378
Qy	7393	aacagcgctcctctgagatgacacaccgcgtcttcgtctcgaagcgggccctcctagcttcgt	7452
Dh	7379	aacagcgctcctctgagatgacacaccgcgtcttcgtctcgaagcgggccctcctagcttcgt	7438
Qy	7453	ctccctgttacttaacagtgtgggccttgcttaatatlaagtgtgtgtcccttgcttagagaa	7512
Dh	7439	ctccctgttacttaacagtgtgggccttgcttaatatlaagtgtgtgtcccttgcttagagaa	7498
Qy	7513	cggagtgaatgcagatccagatcatgtaacttagtgcaacagatlaacaaagcctctcgaaccaa	7572
Dh	7499	cggagtgaatgcagatccagatcatgtaacttagtgcaacagatlaacaaagcctctcgaaccaa	7558
Qy	7573	ggagaaacctgaacctctcagaacctcccaagttcttaagtttagaaactttaacaagaagaag	7632
Dh	7559	ggagaaacctgaacctctcagaacctcccaagttcttaagtttagaaactttaacaagaagaag	7618
Qy	7633	tggggaaatgaaagatagtaaatgtcaaaccttaaacccctccagaacccagaagaattataaaa	7692
Dh	7619	tggggaaatgaaagatagtaaatgtcaaaccttaaacccctccagaacccagaagaattataaaa	7678
Qy	7693	agctcttaaatgcccccgaaattacagacccctgtgcctgcgcgtacgaataatagttagaaggtca	7752
Dh	7679	agctcttaaatgcccccgaaattacagacccctgtgcctgcgcgtacgaataatagttagaaggtca	7738
Qy	7753	caattcctattgtttcacaggcgctgtctatccctgtggcctaagtaagataagaagaatagatt	7812
Dh	7739	caattcctattgtttcacaggcgctgtctatccctgtggcctaagtaagataagaagaatagatt	7798
Qy	7813	gactaatgcgttatctcggagttctgttaaaactgacgtgcgaacataagaagaattgatttaaac	7872
Dh	7799	gactaatgcgttatctcggagttctgttaaaactgacgtgcgaacataagaagaattgatttaaac	7858

QY	7873	agagaacgcccagatgaaccatactcaactcgaaactcgtcacctctgcggccaggccccagc	7932		
XX					
Db	7859	atgtagacgccccaagtaagacctactcaactcgacaactctgtcacctctgcggccaggccccagc	7918		
QY	7933	agatcgagacccctcgcggagcatattttaaatgatatttgctccacgagcgcgggctctcgata	7992		
Db	7919	agttgtagagaccctccgagcatattttaaatgatatttgctccacgagcgcgggctctcgata	7978		
QY	7993	ttttaaataatgatatttgctccacgagcgcgggctctcgataattttaaatgatatttgctg	8052		
Db	7979	ttttaaataatgatatttgctccacgagcgcgggctctcgataattttaaatgatatttgctg	8038		
QY	8053	acgcacaagccttgtgtgtgaacccccataaagcgtctccgattccgcactcgcgggccccga	8112		
Db	8039	acgcacaagccttgtgtgtgaacccccataaagcgtctccgattccgcactcgcgggccccga	8098		
QY	8113	gtcctctacccctcgctgtgtgtgtacgactctgtggccccacgagcgcttggaataaaaatcct	8172		
Db	8099	gtcctctacccctcgctgtgtgtgtgtacgactctgtggccccacgagcgcttggaataaaaatcct	8158		
QY	8173	cttgctgtttgcatcaaaaaaaaaaaaaaaaaaaaaa	8209		
Db	8159	cttgctgtttgcatcaaaaaaaaaaaaaaaaaaaaaa	8195		
<b>RESULT 3</b>					
AAAF77727					
ID	AAAF77727	standard; DNA; 8132 BP.			
XX	AAAF77727;				
AC					
XX					
DT					
XX					
DE	23-MAY-2001	(first entry)			
XX					
Nucleotide sequence of a retrovirus found in miniature swine.					
KW	Retrovirus; graft transplantation; xenotransplantation; miniature swine;				
SS.					
OS					
XX	Unidentified.				
PN	US6190861-B1.				
PD	20-FEB-2001.				
XX					
Pf	13-DEC-1996;	96US-0766528.			
XX					
PR	14-DEC-1995;	95US-0572645.			
XX					
(GEHO ) GEN HOSPITAL CORP.					
Fishman JA;					
DR	WPI; 2001-256211/26.				
DR	P-PSDB; AAB73285, AAB73286, AAB73287.				
Assessing risk of endogenous retroviruses in clinical practice and in					
xenotransplantation, comprises using probe sequences derived from swine					
or miniature swine retroviral genome -					
Claim 1; Fig 3; 127pp; English.					
The present invention relates to a method for screening a cell or tissue					
for the presence or expression of a retrovirus (RV), comprising					
contacting a target nucleic acid from the cell or tissue with a second					
nucleic acid from the present invention (e.g. the present sequence or a					
fragment thereof). The method is useful for RV detection and to assess					
graft transplantation risk. Screening of animals allows the elimination					
of donors with active replication of known viruses. Inactive proviruses					
can be detected and inactivated, allowing identification and elimination					
of potential human pathogens derived from swine in a manner not possible					
in the outbred human organ donor population and is important to the					
development of human xenotransplantation.					



XX Sequence 8132 BP, 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

Query Match 76.5%; Score 6282.4; DB 22; Length 8132;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 7165; Conservative 0; Mismatches 931; Indels 134; Gaps 19;

QY 1 gttgtgtacgactgtgtggcccccagcgctgtgaaataaaatccctctgtctgtgacac 60  
Db 3 gtgtgtacgactgtgtggcccccagcgctgtgaaataaaatccctctgtctgtgacac 62  
QY 61 aagaccgctctcgtgaaatgatt-tgggtgtcgcctctccagaccgagaggggga 119  
Db 63 aagaccgctctcgtgaaatgatttgaaggagtcgcttctccagaccgagag----- 115  
QY 120 ttgtcttctacgtgccttcaattgtgtcgttgccgggaaatccctgtgagaccctct 179  
Db 116 -ttcttctgtcgtccttaacatttgggggtcgtccgggactcgtccggccaccctta 173  
QY 180 aacccgagaaacgactgtgaggtta-aaggatccctctgtgaaacgtgtgtg--tgtgtc 236  
Db 174 aacccgagaaacgactgtgaggttaaaagatccctctttaaagcgtgtatgcatgtacc 233  
QY 237 ggcggcgctcctgtctgtgaatgtctgtcttcgtgtgatgcgccttcggtttgcagctg 296  
Db 234 ggcggcgctcctgtctgtgaatgtctgtcttcgtgtgatgcgccttcggtttgcagctg 293  
QY 297 tcccttcagacgtgaagactgtgaggtatgtatcagacagacgtgtcgtatgagaggttcaag 356  
Db 294 tcccttcagacgtgaagactgtgaggtatgtatcagacagacgtgtcgtatgagaggttcaag 353  
QY 357 gctgcacacgtgtgggagaccccgaggaggtgtgagagagccagggagccgtgtgtgtctcc 416  
Db 354 gctgcacacgtgtgggagaccccgaggaggtgtgagagagccagggagccgtgtgtgtctcc 413  
QY 417 tactgtcgtcagaaggagacggatctctgttttgaagcgaaagctccctcccgcgcg 476  
Db 414 tactgtcgtcagaaggagacggatctctgttttgaagcgaaagctccctcccgcgcg 473  
QY 477 tccgactcttctgctgtctgttgaagacgagcggtcgcgtgtctgtctgtgactcttg 536  
Db 474 tccgactcttctgctgtctgttgaagacgagcggtcgcgtgtctgtctgtgactcttg 533  
QY 537 gttctgtctcgtgtgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 596  
Db 534 gttctgtctcgtgtgtgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtct 593  
QY 597 acagtgtacacccctctgt 656  
Db 594 acagtgtacacccctctgt 653  
QY 657 cataatctgtcagttcaggttaagaaggagacctgtgagacattctctgtgctcgtgaatg 716  
Db 654 cataatctgtcagttcaggttaagaaggagacctgtgagacattctctgtgctcgtgaatg 713  
QY 717 ccaacattcgaatgttgatggcaltcagaggagaccttcaattcgtgaatctatccctgtgt 776  
Db 714 cgaacattcgaatgttgatggcaltcagaggagaccttcaattcgtgaatctatccctgtgt 773  
QY 777 gtttaaggcaatcatttttcaagacgtgacccggctctctacccgtatcagagaccctatac 836  
Db 774 gtttaaggcaatcatttttcaagacgtgacccggctctctacccgtatcagagaccctatac 833  
QY 837 cttaacgtgtgcaagatttggcagaagacccctccgcacatgggtttaacacatgtgcataataa 896  
Db 834 cttaacgtgtgcaagatttggcagaagacccctccgcacatgggtttaacacatgtgcataataa 893  
QY 897 ccaagaaagcaggttcccgaaatccctgtgtcttggagagaaaaaacaacacacgtgcgcga 956  
Db 894 ccaagaaagcaggttcccgaaatccctgtgtcttggagagaaaaaacaacacacgtgcgcga 953  
QY 957 aaagtgcagccctctctctgtatctaccccgagatcagagagcgcgcgactgtgcgcga 1016

Db 954 aaagtgcagccctc-ctctatctaccccgagatgagagacacgcgcttgcgcga 1012  
QY 1017 accccaacgtgttcccccaccccttaccagacaggtgtcgtgtgagggga-cctctgc 1075  
Db 1013 accccaacgtgttcccccaccccttaccagacaggtgtcgtgtgaggggaccccttgc 1072  
QY 1076 cctctcctgtgactcgtgtgtgtgagagacctgtctgcggagactcgtgagcccgagagcg 1135  
Db 1073 cctctcctgtgactcgtgtgtgtgagagacctgtctgcggagactcgtgagcccgagagcg 1132  
QY 1136 caacccgagcgcagacagatcgcataatccgctgtgcacactatgtgacctccat 1195  
Db 1133 caacccgagcgcagacagatcgcataatccgctgtgcacactatgtgacctccat 1192  
QY 1196 gccagggggcacaatttgagagccctccagatcttgcccttctctgcagatctctata 1255  
Db 1193 gccagggggcacaatttgagagccctccagatcttgcccttctctgcagatctctata 1252  
QY 1256 ttggaataactaacatccctcttctcgagagatccccaagcctcagcggtgtgtga 1315  
Db 1253 ttggaataactaacatccctcttctcgagagatccccaagcctcagcggtgtgtga 1312  
QY 1316 gtccctatgttctctcaccagcctactgtggaatgtgtcaacagctgtcgcagacact 1375  
Db 1313 gtccctatgttctctcaccagcctactgtggaatgtgtcaacagctgtcgcagacact 1372  
QY 1376 ctccaacacgagagagcgtgagagaaatctcgttaagggtcctagaaaaatgttctcgggc 1435  
Db 1373 ctccaacacgagagagcgtgagagaaatctcgttaagggtcctagaaaaatgttctcgggc 1432  
QY 1436 cgaacggcgagacccacacgactgtgcaaaatgagatctgacatgtgttccctgtgcgcc 1495  
Db 1433 cgaacggcgagacccacacgactgtgcaaaatgagatctgacatgtgttccctgtgcgcc 1492  
QY 1496 cgggttgggactacacacggtcgaaggtagggagagcttgaataatctatccagagctct 1555  
Db 1493 cgggttgggactacacacggtcgaaggtagggagagcttgaataatctatccagagctct 1552  
QY 1556 ggttcggcggtctcgcggcgccctcgaagacgcccataatttggctcaagttgaagaggt 1615  
Db 1553 ggttcggcggtctcgcggcgccctcgaagacgcccataatttggctcaagttgaagaggt 1612  
QY 1616 gatgcagggagacgaaacacccctcgcgttatctctgtgagagcctatgtgaagccttcag 1675  
Db 1613 gatgcagggagacgaaacacccctcgcgttatctctgtgagagcctatgtgaagccttcag 1672  
QY 1676 gcggttcaaccccttctgtatcctacacgagcagagccagaagcctcagtgccctgcct 1735  
Db 1673 gcggttcaaccccttctgtatcctacacgagcagagccagaagcctcagtgccctgcct 1732  
QY 1736 catgtgcagctgcgtctgtgatatcaggaagaacttcgaagaactgtgaagggtttcaagga 1795  
Db 1733 catgtgcagctgcgtctgtgatatcaggaagaacttcgaagaactgtgaagggtttcaagga 1792  
QY 1796 ggtcgtgttctacgtctcgtatgagagagcagaagaggtgttcaagaaggagagacaa 1855  
Db 1793 ggtcgtgttctacgtctcgtatgagagagcagaagaggtgttcaagaaggagagacaa 1852  
QY 1856 agagggagagagagacagagaaagaaagagagagagagagagagagagagagagagag 1915  
Db 1853 agagggagagagagacagagaaagaaagagagagagagagagagagagagagagagag 1912  
QY 1916 acggcagaagagagaaatttgaactaagaatctgtgcgagaggtgttgaaggagagagagag 1975  
Db 1913 acggcagaagagagaaatttgaactaagaatctgtgcgagaggtgttgaaggagagagagag 1972  
QY 1976 ggaagagagagagagatttgaagaaatagtcacagcccttagacagctcaggaagacctggg 2035  
Db 1973 ggaagagagagagagatttgaagaaatagtcacagcccttagacagctcaggaagacctggg 2032  
QY 2036 catatggacccactcgaagagacagaggtgtgtctatgttaagaaaaagacactggcg 2095

Db 2033 caataaggaccccaactcgacaaagcaatgltgacatattgtaaaagaagagacatggtgc 2092  
QY 2096 aaggactgcccccaagaagagaaacaaaggaacgaaggtccttaagcttagaagaataa 2155  
Db 2093 aaggaaactgcccccaagaagagaaacaaaggaacgaaggtccttaagcttagaagaataa 2152  
QY 2156 agattaaaggagacggtgttcggaacccctcccgagccagggtaactttgaagtgtga 2215  
Db 2153 agattaaaggagacggtgttcggaacccctcccgagccagggtaactttgaagtgtga 2212  
QY 2216 ggggcaacccattgaattcctggttgaatacggagcggaacattcagttgtgtcacacac 2275  
Db 2213 ggggcaacccattgaattcctggttgaatacggagcggaacattcagttgtgtcacacac 2272  
QY 2276 attggaataaataagaataaaataatcctggttgaatacggagcggaacattcagttgtgtcacacac 2335  
Db 2273 attggaataaataagaataaaataatcctggttgaatacggagcggaacattcagttgtgtcacacac 2332  
QY 2336 tccatggaactacccgaagaacggttgaacttggaggtggagcggttaacccactcgtttct 2395  
Db 2333 tccatggaactacccgaagaacggttgaacttggaggtggagcggttaacccactcgtttct 2392  
QY 2396 ggtatccctgaagtgcccaagatccctcttaaggtagaagacttaactgcaagaagatggagac 2455  
Db 2393 ggtatccctgaagtgcccaagatccctcttaaggtagaagacttaactgcaagaagatggagac 2452  
QY 2456 tcaaatctcttgaacaaaggaagaaagagatgtctgtgaataacaaacccactcaactgt 2515  
Db 2453 acaaatctcttgaacaaaggaagaaagagatgtctgtgaataacaaacccactcaactgt 2512  
QY 2516 gttgaacctccaattgaatgataatgataatgataatctccccaagttaagctgtatca 2575  
Db 2513 gttgaacctccaattgaatgataatgataatgataatctccccaagttaagctgtatca 2572  
QY 2576 agatatacagtcctggttggagcagtttccccaagccttggcgagaaacccgaagatggag 2635  
Db 2573 aatataataatctcgttggagacagtttccccaagccttggcgagaaacccgaagatggag 2632  
QY 2636 ttggcgaaagcaagttcccccacaggtatcaactgaagggccagtgctacacacagatc 2695  
Db 2633 ttggcgaaagcaagttcccccacaggtatcaactgaagggccagtgctacacacagatc 2692  
QY 2696 agtaagacagatcccttggatagagaggtcgcgaagaagatttggcgagatgttcaag 2755  
Db 2693 agtaagacagatcccttggatagagaggtcgcgaagaagatttggcgagatgttcaag 2752  
QY 2756 attaatccaacagggatcctagttcctgttcccaatcccttggatatacctccctgtacac 2815  
Db 2753 attaatccaacagggatcctagttcctgttcccaatcccttggatatacctccctgtacac 2812  
QY 2816 ggttaagaagccttggagacaaatgataatcgacacagtaacaggaacttgaagaggtcaataa 2875  
Db 2813 ggttaagaagccttggagacaaatgataatcgacacagtaacaggaacttgaagaggtcaataa 2872  
QY 2876 aagggttcagagacatacacaacaaaggtcccgaaacctataactcttggcgccctcc 2935  
Db 2873 aagggttcagagacatacacaacaaaggtcccgaaacctataactcttggcgccctcc 2932  
QY 2936 ggcgtgaacggaacttgcacacagatattggaacttaagaatgaccttcttcgctgtgatt 2995  
Db 2933 accccaacggaacttgcacacagatattggaacttaagaatgaccttcttcgctgtgatt 2992  
QY 2996 acaccaactagccaacacacttttgccttcgaatgagagatccaggttaccggaagaac 3052  
Db 2993 acaccaactagccaacacacttttgccttcgaatgagagatccaggttaccggaagaac 3052  
QY 3056 cgggcaactacacttgcgaacccgagatccccaaggttcaaaaactcccgacatctttga 3115  
Db 3053 cgggcaactacacttgcgaacccgagatccccaaggttcaaaaactcccgacatctttga 3112  
QY 3116 cgaagccctacacagagagacttgcgaacttcagatccaacaccccttaagttgacccctc 3175  
Db 3113 cgaagccctacacagagagacttgcgaacttcagatccaacaccccttaagttgacccctc 3172

QY 3176 ccagtaactggaatgaccttcttcgcggaagccaccaacagagactttagaaggtac 3235  
Db 3173 ccagtaactggaatgaccttcttcgcggaagccaccaacagagactttagaaggtac 3232  
QY 3236 gaagcactactgtctggaattgtctgacatagatcagagaccttctgtaagaagccca 3295  
Db 3233 gaagcactactgtctggaattgtctgacatagatcagagaccttctgtaagaagccca 3292  
QY 3296 gatttcaggaagagagatcacttgggtgacagtttggcgggcggaacgataagct 3355  
Db 3293 gatttcaggaagagagatcacttgggtgacagtttggcgggcggaacgataagct 3352  
QY 3356 gacggagcgacaggaagaaactgtagttccagataccggtcccaacacagccaacagat 3415  
Db 3353 gacggagcgacaggaagaaactgtagttccagataccggtcccaacacagccaacagat 3412  
QY 3416 gaagaggttttggggacagcctggaatttgcagacttggatcccggttggagactt 3475  
Db 3413 gaagaggttttggggacagcctggaatttgcagacttggatcccggttggagactt 3472  
QY 3476 agcagcccaactctacccgtacacaaagaaagggagatctctctggtcctcgagca 3535  
Db 3473 agcagcccaactctacccgtacacaaagaaagggagatctctctggtcctcgagca 3532  
QY 3536 ccagaaggaatttgaatgatacaaaaagccctgtgagcgacactgtctggtccctcc 3595  
Db 3533 ccagaaggaatttgaatgatacaaaaagccctgtgagcgacactgtctggtccctcc 3592  
QY 3596 ttgagtaacttaaccccttacccttattgttgaatgagcgtaagggagtagcccggaaggt 3655  
Db 3593 ttgagtaacttaaccccttacccttattgttgaatgagcgtaagggagtagcccggaaggt 3652  
QY 3656 tttaaccacaaccccttagagacccaatgagagacacttgcctacactgttcaaaaagctga 3715  
Db 3653 tttaaccacaaccccttagagacccaatgagagacacttgcctacactgttcaaaaagctga 3712  
QY 3716 tctctgaagcagttgttggccgttatgtctgaaggtatcgagcttgggtccatctgt 3775  
Db 3713 tctctgaagcagttgttggccgttatgtctgaaggtatcgagcttgggtccatctgt 3772  
QY 3776 caagagcgtctgaacaaattgaacttggagacagataatgaatgaagccccaatga 3835  
Db 3773 caagagcgtctgaacaaattgaacttggagacagataatgaatgaagccccaatga 3832  
QY 3836 ggaagaacatcgttctggcagcccccagacagatgataatgaacaaagcccgatgacacata 3895  
Db 3833 ggaagaacatcgttctggcagcccccagacagatgataatgaacaaagcccgatgacacata 3892  
QY 3896 tcaagcctgtcttcaacagagaaggtcaacttgcctcaacagcgcgtctcaaaccttgc 3955  
Db 3893 tcaagcctgtcttcaacagagaaggtcaacttgcctcaacagcgcgtctcaaaccttgc 3952  
QY 3956 caactcttgccttgaaggaacatgataacaaatgataatgataatgataatgataat 4015  
Db 3953 caactcttgccttgaaggaacatgataacaaatgataatgataatgataatgataat 4012  
QY 4016 tgaagagacttgggttccgaagaagacttcaacacataccgctgagatggaaggtgctaac 4075  
Db 4013 tgaagagacttgggttccgaagaagacttcaacacataccgctgagatggaaggtgctaac 4072  
QY 4076 ctggttcaactgaacggaacagctatgtgttgaaggttaagagatgtgtggcgagct 4135  
Db 4073 ctggttcaactgaacggaacagctatgtgttgaaggttaagagatgtgtggcgagct 4132  
QY 4136 ggttgaagcggagcccgagacatcttggcgagcagcgttgcgggaaggaacttcaagcaca 4195  
Db 4133 ggttgaagcggagcccgagacatcttggcgagcagcgttgcgggaaggaacttcaagcaca 4192  
QY 4196 ggtcagatcatalggtccctacgcaagcttggcggttggcggaaggaatacctaatacat 4255  
Db 4193 ggtcagatcatalggtccctacgcaagcttggcggttggcggaaggaatacctaatacat 4252

```
QY 4256 ttatacgacagcaggtatgctcttgacactgcacacgttacacgggcacatctaaca 4315
Db 4253 ttatacgacagcaggtatgctcttgacactgcacacgttacatggggccatctaaca 4312
QY 4316 aagggggttgcttaacctcagcagggagggaataaagaacaagaagaaattctaagcct 4375
Db 4313 aagggggttgcttaacctcagcagggagggaataaagaacaagaagaaattctaagcct 4372
QY 4376 attgaagaccttaattggccaaaaagctagctataatacactgctccctggacataa 4435
Db 4373 attgaagaccttaattggccaaaaagctagctataatacactgctccctggacataa 4432
QY 4436 agccaagaatctcatatctagagggaacacagatgctgcacgggttgccagacggagc 4495
Db 4433 agctaaagatctcatatctcagagggaacacagatgctgcacgggttgccagacggagc 4492
QY 4496 ccagggttgcttaacctcctcctaaatagaacacgccaagccccaagaacccagacgaca 4555
Db 4493 ccagggttgcttaacctcctcctaaatagaacacgccaagccccaagaacccagacgaca 4552
QY 4556 gtacacacctagaagactgycgacagagataaaaaagataagacagatctctctagactccgga 4615
Db 4553 gtacacacctagaagactgycgacagagataaaaaagataagacagatctctctagactccgga 4612
QY 4616 ggggacctgtcatacctcatatgtgggaaggaatcctgcgccacaagaagaggttagata 4675
Db 4613 agggacctgtcatacctcatatgtgggaaggaatcctgcgccacaagaagaggttagata 4672
QY 4676 tgtccacaagatacatcgtcttaacccacactaggaactaaacaccttgacagatgtgcag 4735
Db 4673 tgtccacaagatacatcgtcttaacccacactaggaactaaacaccttgacagatgtgcag 4732
QY 4736 aacatccoctatataatgttctcgaagctacacagagatgctgcagctcggtgtgcaaacatg 4795
Db 4733 aacatccoctatataatgttctcgaagctacacagagatgctgcagctcggtgtgcaaacatg 4792
QY 4796 tgtccoctgcagctgtgttaattgttaactcctccagaaatccctccagaaagagactaag 4855
Db 4793 tgtccoctgcagctgtgttaattgttaactcctccagaaatccctccagaaagagactaag 4852
QY 4856 gggagacccacccagcgcctcaactgycgagatgagactcaactgaagttaaagccgcgtataa 4915
Db 4853 gggagacccacccagcgcctcaactgycgagatgagactcaactgaagttaaagccgcgtataa 4912
QY 4916 cggaaaaaataatattatgttcttctgaagacacttctcagatagggttaagcgtctatcc 4975
Db 4913 cggaaaaaataatattatgttcttctgaagacacttctcagatagggttaagcgtctatcc 4972
QY 4976 tactaagaagaagacttcaacccgtgctggtctaaagaataactctggaagaatcttctcaag 5035
Db 4973 tactaagaagaagacttcaacccgtgctggtctaaagaataactctggaagaatcttctcaag 5032
QY 5036 atttggaaatacctaagtaataagggtcagacaatgctccagcttctgctgcccagtaag 5095
Db 5033 atttggaaatacctaagtaataagggtcagacaatgctccagcttctgctgcccagtaag 5092
QY 5096 tcaaggacactggccaaagattgggattgtattggaactcatgattgctacatacagaccaca 5155
Db 5093 tcaaggacactggccaaagattgggattgtattggaactcatgattgctacatacagaccaca 5152
QY 5156 aagctcaggaacagtagaagaagatgaataagacacatlaaagaagaccccttaccaatgtac 5215
Db 5153 aagctcaggaacagtagaagaagatgaataagacacatlaaagaagaccccttaccaatgtac 5212
QY 5216 cagcagagactggaacttaattgattgattgctcctcctgcttctttagaggttag 5275
Db 5213 cagcagagactggaacttaattgattgattgctcctcctgcttctttagaggttag 5272
QY 5276 gaacaccccttgacagattgctgaccccttaagtaattgcttcaaggggagacccccc 5335
Db 5273 gaacaccccttgacagattgctgaccccttaagtaattgcttcaaggggagacccccc 5332
QY 5336 gttagcagaataatgctcttgacacatagtgctgattgctcttccagcacttgctc 5395
Db 5333 attgttagaataattgctcttgacacatagtgctgacgctgctcttccagcacttgctc 5392
QY 5396 taagctcaagcgctcagtgagtgagcagcagcgtggaagcagctccggagagcta 5455
Db 5393 taagctcaagcgctcagtgagtgagcagcagcgtggaagcagctccggagagcta 5452
QY 5456 ctc---aggaaggacttgcaagttccacatcgctccgaagtctgagctcagtgat 5512
Db 5453 ctcaggagagaggacttgcaagttccacatcgctccgaagtctgagctcagtgatcagtgat 5512
QY 5513 taagcgcacacgctgcaggaacacctcagagactcgttggaaggagacacttactcgtacttt 5572
Db 5513 taagcgcacacgctgcaggaacacctcagagactcgttggaaggagacacttactcgtacttt 5572
QY 5573 gaccacacaaacgctgtgaaagtcgaagaagatccaccactgatalcatgcatccagct 5632
Db 5573 gaccacacaaacgctgtgaaagtcgaagaagatccaccactgatalcatgcatccagct 5632
QY 5633 taagccggcgccacccctccgactcgggtgagagagccgaagaagactggaatccoctta 5692
Db 5633 taagccggcgccacccctccgactcgggtgagagagccgaagaagactggaatccoctta 5692
QY 5693 gcttcgctccacacgctgtgcttctactctcaacaataactccacgagccagtagtaaac 5752
Db 5693 gcttcgctccacacgctgtgcttctactctcaacaataactccacgagccagtagtaaac 5752
QY 5753 gacctatagaacagctcgaacccccaatagaaccttactccctacactgctgattatgacc 5812
Db 5753 gacctatagaacagctcgaacccccaatagaaccttactccctacactgctgattatgacc 5812
QY 5813 ctgatacgggtgtcactctgtaataagactcagaggtgtgctcctagaggcactgtgac 5872
Db 5813 ctgatacgggtgtcactctgtaataagactcagaggtgtgctcctagaggcactgtgac 5872
QY 5873 ctgacatgaattctctcctccgactgatacttaaccccgctttaaagaacacactccaca 5932
Db 5873 ctgacatgaattctctcctccgactgatacttaaccccgctttaaagaacacactccaca 5932
QY 5933 ctgactatagaagtttgccacagatcagatgagtagagtttctctatgcaacactataca 5929
Db 5933 ctgactatagaagtttgccacagatcagatgagtagagtttctctatgcaacactataca 5929
QY 5933 taagccgtatgtaggtgtcctatctgctgcacagacagc---agaaagaataactgtg 5989
Db 5933 taagccgtatgtaggtgtcctatctgctgcacagacagc---agaaagaataactgtg 5989
QY 5990 ggggtctctggaatccctctctgtagagagatgagcgtgcctccacacagatgagact 6049
Db 5990 ggggtctctggaatccctctctctgtagagagatgagcgtgcctccacacagatgagact 6049
QY 5990 gaaatccagagatcttctcttgtaaacaatggaactgtaaccttcaatgtagatatt 6049
Db 5990 gaaatccagagatcttctcttgtaaacaatggaactgtaaccttcaatgtagatatt 6049
QY 6050 ggaatgagcagatctctccagagacgggtlaaatacttctccttgca----- 6098
Db 6050 ggaatgagcagatctctccagagacgggtlaaatacttctccttgca----- 6098
QY 6050 ggaatgagcagatctctccagagacgggtlaaatacttctccttgca----- 6098
Db 6050 ggaatgagcagatctctccagagacgggtlaaatacttctccttgca----- 6098
QY 6099 -ttccggccggcgaagtacaaaatgaatgaactatataagaataagactgctccocat 6157
Db 6099 -ttccggccggcgaagtacaaaatgaatgaactatataagaataagactgctccocat 6157
QY 6110 gctctgtaacatlaattacttactgactgactgagactlgaagccccaagtgtctcctt 6169
Db 6110 gctctgtaacatlaattacttactgactgactgagactlgaagccccaagtgtctcctt 6169
QY 6158 cagacttagatlatctaaagaataagtttact--gaaaggaacaacaggaataattcaaa 6215
Db 6158 cagacttagatlatctaaagaataagtttact--gaaaggaacaacaggaataattcaaa 6215
QY 6170 cagacccagatlaacccaataaataagtttcaactgagagaagaagaacaagaataatccaa 6229
Db 6170 cagacccagatlaacccaataaataagtttcaactgagagaagaagaagaataatccaa 6229
QY 6216 agtgaataatgtagaactcgtgggaatagtttattatataatgycggg---gagacag 6271
Db 6216 agtgaataatgtagaactcgtgggaatagtttattatataatgycggg---gagacag 6271
QY 6230 aatgggttaattgtagtcttgctggggaatgtagatattatgaggtcgtgggtaacaacag 6289
Db 6230 aatgggttaattgtagtcttgctggggaatgtagatattatgaggtcgtgggtaacaacag 6289
QY 6272 ggtccacttaacacactgcttaggataagacggggaacacccctgtgcaatg 6331
Db 6272 ggtccacttaacacactgcttaggataagacggggaacacccctgtgcaatg 6331
QY 6290 gctccacttaacatctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 6346
Db 6290 gctccacttaacatctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 6346
QY 6332 gaccgataaagactgagctgagacaggggcccggcccttgagacacccgataactgac 6391
Db 6332 gaccgataaagactgagctgagacaggggcccggcccttgagacacccgataactgac 6391
QY 6347 gaccataatcgtctgagcgggtcaagaacccccaaccc----- 6385
Db 6347 gaccataatcgtctgagcgggtcaagaacccccaaccc----- 6385
QY 6392 cgtgcccacatlaacactgctgctgagcctgacataacacagcgcctgagcaacagtag 6451
Db 6392 cgtgcccacatlaacactgctgctgagcctgagcctgacataacacagcgcctgagcaacagtag 6451
```

[illegible]

QY	7532	tcacgtgacttaaggcaacagagaccagagccctctcgagccaagaggaatggacctacgc	7591
Db	7481	tcaatgtacttaagcaacagatccaaagccgcgtctcagc---agggaagctggcgcgtacgc	7533
QY	7592	ctcccacttctaagattgaactattaaacaagacagaagctggggaaatgaaagagatga	7651
Db	7538	tctaccagcttctaagattgaactattaaacaagagagaagctggggaaatgaaagagatga	7597
QY	7652	aatgcacacttaaccctcccaagaaccagaagtttaataaagcttaaatgcgcccgaa	7711
Db	7598	aatacaaaccttaagctaaatgagaagctttaaataatgtcttgaaatccagagttgttcccta	7657
QY	7712	ttacagaccctgcgcgtgcacagaatagtgagaaggtccacatctccattgttccagg	7771
Db	7658	tagttaaaaagattagttttttttgtctgttttaaatatgtcggaagttaaatagccctag	7717
QY	7772	gcctgctacatccctggcct-aagtaagataacagagaatagtgtagc---taatcgctat	7826
Db	7718	tacatgctctctagcagctgaactctcttgaaactatttgagatatacaagaagaagagattc	7777
QY	7827	ctggaattctgtaaaactgactgtgcacatagaagaattgattacacattgacagccctag	7886
Db	7778	ctaaactgtctgttttagctctctgtaaaactgtgtcgcctaaagaatgttgaaatgtgat	7837
QY	7887	tgaacctactcaactgcaatctgtcaactgtgccagagaaccgaagatgvcgagccctcc	7946
Db	7838	acaactatcctgtgtgacacaacatgtctccaccaccgccgaacatgtcgcaaatgtgtaactc	7897
QY	7947	ggagctattttaaaatgatatgtgtccacggaacgvcggtctcgcgataattttaaatgattg	8006
Db	7898	taaaacaattttaaatattatgtgtccacgaagcgcggtctctcgaaagttttaaatgtactg	7957
QY	8007	gtccatggaagcgcggtctcgcgataattttaaatagttgtgttgagacgcaagcttg	8066
Db	7958	gtttgt-----gatatttgaatgattgtgttgtaaagcgcggtcttg	8002
QY	8067	ttgtgaaccccatlaaaagcgtgtccgaattccgcacatcgvgcgcgcagctctctacacctg	8126
Db	8003	ttgtgaaccccatlaaaagcgtgtccgaattccgcacatcgvgcgcgcagctctctacacctg	8063
QY	8127	cgtagtgaactgctgtggccccaagcgcgttggaataaaatccctctgtctgtttgat	8186
Db	8063	cgtagtgaactgctgtggccccaagcgcgttggaataaaatccctctgtctgtttgat	8122
QY	8187	caaaaaaaa 8196	
Db	8123	caaaaaaaa 8132	
RESULT 4			
AAT74884			
ID	AAT74884	standard; cDNA; 7892 bp.	
XX	AAT74884;		
AC			
XX			
DT	09-FEB-1998	(first entry)	
XX			
DE	Miniature swine retrovirus cDNA.		
XX			
KW	Retrovirus; porcine; GAG protein; POL protein; ENV protein;		
KW	xenotransplantation; infectious; provirus; organ transplant; donor;		
KW	activated virus; PCR; ss.		
XX			
OS	Porcine retrovirus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	585..2159	
FT		/*lag= a	
FT	mat_peptide	585..2156	
FT		/*tag= b	
FT		/note= "putative GAG protein"	
FT	CDS	2307..5744	

FT mat-peptide /tag= C  
 FT 2307..5741 /tag= d  
 FT /note= "putative POL protein"  
 FT CDS 5620..7536  
 FT /tag= e  
 FT mat-peptide 5620..7533  
 FT /tag= f  
 FT /note= "putative ENV protein"  
 PN MO9721836-A1.  
 PD 19-JUN-1997.  
 PF 13-DEC-1996; 96WO-US19680.  
 PR 14-DEC-1995; 95US-0572645.  
 XX (GHEO ) GEN HOSPITAL CORP.  
 PA Fishman JA;  
 PI MPI: 1997-332804/30.  
 DR P-PSDB; AAM32096-W32098.  
 XX  
 PT New nucleic acid from porcine retroviruses - used for detecting  
 PT viruses in transplant or other tissue and for assessing risk of  
 PT transmitting infection to graft recipient  
 XX  
 PS Claim 22; Fig 3; 128pp; English.  
 CC This CDNA sequence represents a porcine retrovirus from miniature swine  
 CC containing the putative coding regions for viral GAG, POL and ENV  
 CC proteins. This sequence and PCR fragments generated from it  
 CC (see AAT74812-T74882) can be used to screen organs for the presence of  
 CC porcine retroviruses prior to xenotransplantation. Transplantation can  
 CC increase the likelihood of retroviral activation if intact and  
 CC infectious proviruses are present. The porcine retroviral sequence can be  
 CC used to generate probes to determine the level (e.g. copy number) of  
 CC intact (i.e. potentially replicating) porcine provirus sequences in a  
 CC strain of xenograft transplantation donors. It can be used to detect  
 CC mutations, genetic lesions or viral recombinants and also to determine  
 CC the histological localisation of activated retroviruses. Using Polymerase  
 CC Chain Reaction DNA Quantitation (PQ) on blood mononuclear cells,  
 CC infectivity titration and susceptibility testing can be performed.  
 CC Intactively animal donors without intact porcine retroviral sequences or a  
 CC lower copy number of viral elements could be selected.  
 CC  
 CC Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;  
 SO  
 Query Match 74.8%; Score 6141; DB 18; Length 7892;  
 Best Local Similarity 89.0%; Pred. No. 0;  
 Matches 6856; Conservative 0; Mismatches 735; Indels 114; Gaps 16;

Db 234 ggcgcgctctctgtctcagtgctgtttccagttgtgcgcgcttccgtttccagctg 293  
 QY 297 tccctcagacgcgtaaaagacttgagagcttgatccagcaagctgctcagaagatacaag 356  
 Db 294 tccctcagacgcgtaaaagacttgagagcttgatccagcaagctgctcagaagatacaag 353  
 QY 357 gctgcaccccttgaggagaccccgaggagtgaggagagcagagagacgctgctgtctcc 416  
 Db 354 gctgctgccccttgaggagaccccgaggagtgaggagagcagagagacgctgctgtctcc 413  
 QY 417 tactctgtcgaagagacgagatctgtgtgttgaagcaagcttccctccgcgcgcg 476  
 Db 414 tactctgtcgaagagacgagatctgtgtgttgaagcaagcttccctccgcgcgcg 473  
 QY 477 tccgaccttttcgctgctgttgaaagacgcagagagtgctgtgtgtgtgtgtgtgt 556  
 Db 474 tccgaccttttcgctgctgttgaaagacgcagagagtgctgtgtgtgtgtgtgtgt 553  
 QY 537 gttctgtctcgtgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596  
 Db 534 gttctgtctcgtgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 593  
 QY 597 acagtgactacccccccttaagttgaactctcagaccattgagctgaagtlagaccaggt 656  
 Db 594 acggtgacgacccccccttaagttgaactctcagaccattgagctgaagtlagaccaggt 653  
 QY 657 cataattgtcagttgaagttgaagagagagagagagagagagagagagagagagagag 716  
 Db 654 cataattgtcagttgaagttgaagagagagagagagagagagagagagagagagagag 713  
 QY 717 ccaacattcagattgtgattgacatcagagagagagagagagagagagagagagagagag 776  
 Db 714 cgcagattcagattgtgattgacatcagagagagagagagagagagagagagagagagag 773  
 QY 777 gttaaagcaatcatttttaagacttgacccgcgtctcattccatcagagagagagagag 836  
 Db 774 gttaaagcaatcatttttaagacttgacccgcgtctcattccatcagagagagagagag 833  
 QY 837 cttaagtggaagatttgag 896  
 Db 834 cttaagtggaagatttgag 893  
 QY 897 ccaagaaagcaggttcccggaatccctgtgtctcttgagagagagagagagagagagagag 956  
 Db 894 ccaagaaagcaggttcccggaatccctgtgtgtctcttgagagagagagagagagagagagag 953  
 QY 957 aaagtgcagccctctctctgtatcaccgcgagatcagagagagagagagagagagagag 1016  
 Db 954 aaagtgcagccctctctctgtatcaccgcgagatcagagagagagagagagagagagag 1012  
 QY 1017 accccaacctgttcccccaccccttatccagacagaggtgtgtgtgtgtgtgtgtgtgt 1075  
 Db 1013 accccaacctgttcccccaccccttatccagacagaggtgtgtgtgtgtgtgtgtgtgt 1072  
 QY 1076 ccctcttgagacctcgt 1135  
 Db 1073 ccctcttgagacctcgt 1132  
 QY 1136 caccgcgagcagagacagagagatcgcgatatcaccgtctgcgacctatgtgcccctccat 1195  
 Db 1133 caccgcgagcagagacagagagatcgcgatatcaccgtctgcgacctatgtgcccctccac 1192  
 QY 1196 gccaggaggcccaattgcagccctcagagatgtgacctttctctcagagactctatata 1255  
 Db 1193 gccaggaggcccaattgcagccctcagagatgtgacctttctctcagagactctatata 1252  
 QY 1256 ttggaacctaaacatccccccttctcggagagatccccaagcgtccacaggggttggtgga 1315  
 Db 1253 ttggaacctaaacatccccccttctcggagagatccccaagcgtccacaggggttggtgga 1312  
 QY 1316 gtccctatgttctcaccagcactacttgagatgtatgttaacaagctgtcagagact 1375  
 Db 1313 gtccctatgttctcaccagcactacttgagatgtatgttaacaagctgtcagagact 1372

QY 1376 cttcaaacgagagcgagagagaattctgttagagctagaanaaaatgttccctgggc 1435  
Db 1373 cttcaaacgagagcgagagagaattcttagagctagaanaaaatgttccctgggc 1432  
QY 1436 cgacggcgacacccgagctgtgcaaaaatgagatgacatggatttcccttgatccgc 1495  
Db 1433 cgacggcgacacccgagctgtgcaaaaatgagatgacatggatttcccttgatccgc 1492  
QY 1496 cggtttggactacaacacgctgtagagtgagagagcttgaataatctatcgccagctct 1555  
Db 1493 cggtttggactacaacacgctgtagagtgagagagcttgaataatctatcgccagctct 1552  
QY 1556 ggttggcggtctccggggcgctcctaagaacggccactaatttggcttaagtgaaagagct 1615  
Db 1553 ggttggcggtctccggggcgctcctaagaacggccactaatttggcttaagtgaaagagct 1612  
QY 1616 gatcagagacgagacgaacctccctcgtaatttctttagagagctcaltgaaagccttcag 1675  
Db 1613 gatcagagacgagacgaacctccctcgtaatttctttagagagagctcaltgaaagccttcag 1672  
QY 1676 gcggttcaaccttttgatcctactacagagagcccaagaagcctcagtgccctggcctt 1735  
Db 1673 gcggttcaaccttttgatcctactacagagagcccaagaagcctcagtgccctggcctt 1732  
QY 1736 catggcgagctcggtcttgatcaltcagaagaacattcagaagacttgaaaggttacaaga 1795  
Db 1733 catggcgagctcggtcttgatcaltcagaagaacattcagaagacttgaaaggttacaaga 1792  
QY 1796 ggtctagttacgtgaactagtgagagagcgagagagtgatatacgaagggagacaga 1855  
Db 1793 ggtctagttacgtgaactagtgagagagcgagagagtgatatacgaagggagacaga 1852  
QY 1856 ag 1915  
Db 1853 ag 1912  
QY 1916 acggaag 1975  
Db 1913 acggaag 1972  
QY 1976 ggaag 2035  
Db 1973 ggaag 2032  
QY 2036 caatag 2095  
Db 2033 caatag 2092  
QY 2096 aaggaactgcgcccaag 2155  
Db 2093 aaggaactgcgcccaag 2152  
QY 2156 agattag 2215  
Db 2153 agattag 2212  
QY 2216 ggggaacacagcttgatctccgtgtgtatcagcgagcgagagagagagagagagagagag 2275  
Db 2213 ggggaacacagcttgatctccgtgtgtatcagcgagcgagagagagagagagagagagag 2272  
QY 2276 attaggaac 2335  
Db 2273 attaggaac 2332  
QY 2336 tccatag 2395  
Db 2333 tccatag 2392  
QY 2396 ggtacccctgagtgccacagtaacctctagtgtagaagacttactgacacagatggagagc 2455  
Db 2393 ggtacccctgagtgccacagtaacctctagtgtagaagacttactgacacagatggagagc 2452  
QY 2456 tcaattctcttgaaacaagaagagacagagagtgctgtgtaatacaaacacacacacacac 2515  
Db 2453 tcaattctcttgaaacaagaagagacagagagtgctgtgtaatacaaacacacacacacac 2512  
QY 2516 gttgacccctccaaatagatgataatataatgactataatctcccaagtgaaagcctgata 2575  
Db 2513 gttgacccctccaaatagatgataatataatgactataatctcccaagtgaaagcctgata 2572  
QY 2576 agatataatgctctgtgttgagagagtggtcccaagccttgagagagagagagagagagag 2635  
Db 2573 agatataatgctctgtgttgagagagtggtcccaagccttgagagagagagagagagagag 2632  
QY 2636 ttctgcaaaagcaggttcccccacaggttatacacttgaaagcgagtgctacacagatc 2695  
Db 2633 ttctgcaaaagcaggttcccccacaggttatacacttgaaagcgagtgctacacagatc 2692  
QY 2696 agtcaagacagtgaccccttgatgaaagagacagagagagagagagagagagagagagag 2755  
Db 2693 agtcaagacagtgaccccttgatgaaagagacagagagagagagagagagagagagagag 2752  
QY 2756 attaatccaaagcaggtctctgttccatccctcttggaatccctctgacacacacacacac 2815  
Db 2753 attaatccaaagcaggtctctgttccatccctcttggaatccctctgacacacacacacac 2812  
QY 2816 ggttagaagccttgagacacatgatatcagacagtgacagagcttgagagagagagagagag 2875  
Db 2813 ggttagaagccttgagacacatgatatcagacagtgacagagcttgagagagagagagagag 2872  
QY 2876 aaggtgtcag 2935  
Db 2873 aaggtgtcag 2932  
QY 2936 gcttcacag 2995  
Db 2933 gcttcacag 2992  
QY 2996 acaccccaactagcacaac 3055  
Db 2993 acaccccaactagcacaac 3052  
QY 3056 cgggagagctac 3115  
Db 3053 cgggagagctac 3112  
QY 3116 cgaagccctacacag 3175  
Db 3113 cgaagccctacacag 3172  
QY 3176 ccaagtacgtgagatgacctctctctgagagagagagagagagagagagagagagagag 3235  
Db 3173 ccaagtacgtgagatgacctctctctgagagagagagagagagagagagagagagagag 3232  
QY 3236 gaagagcactacgtctggaattgtctgacacacacacacacacacacacacacacacacac 3295  
Db 3233 gaagagcactacgtctggaattgtctgacacacacacacacacacacacacacacacacac 3292  
QY 3296 gatttgcag 3355  
Db 3293 gatttgcag 3352  
QY 3356 gacggag 3415  
Db 3353 gacggag 3412  
QY 3416 gagaagatttttggagagacagctgattttgacagagcttgatcccgagggtttggagacct 3475  
Db 3413 gagaagatttttggagagacagctgattttgacagagcttgatcccgagggtttggagacct 3472  
QY 3476 agcagcccaactacacagctac 3535  
Db 3473 agcagcccaactacacagctac 3532  
QY 3536 ccag 3595

Db 3533 ccggaagacatttgatgcatcaaaaagccctgacgacgacccctgctcc 3592  
Qy 3596 tgaagtaactaaaccccttacccttatgtgtagagcgtaaggagtagcccgagagt 3655  
Db 3593 tgaagtaactaaaccccttacccttatgtgtagagcgtaaggagtagcccgagagt 3652  
Qy 3656 tttaaccacaaccccttagagacattgagagagacctgtgtcctaccctgtcaaaagacttga 3715  
Db 3653 tttaaccacaaccccttagagacattgagagagacctgtgtcctaccctgtcaaaagacttga 3712  
Qy 3716 tccctgagccagttgtgtgcccatactgtcgtgaagactatcgacgtgtgcccatactgtgt 3775  
Db 3713 tccctgagccagttgtgtgcccatactgtcgtgaagactatcgacgtgtgcccatactgtgt 3772  
Qy 3776 caagagccgtgacaatttacttctgagacagaataaactgttaataagccccctacgt 3835  
Db 3773 caagagccgtgacaatttacttctgagacagaataaactgttaataagccccctacgt 3832  
Qy 3836 ggaagaacatcgtltcgacagccccacagacgatgga tgaacaaacgcccagatgacccacta 3895  
Db 3833 ggaagaacatcgtltcgacagccccacagacgatgga tgaacaaacgcccagatgacccacta 3892  
Qy 3896 tcaaaagccgtcttccacaagagaggttaacttctgctccacacgcccgtctcaacccctgc 3955  
Db 3893 tcaaaagccgtcttccacaagagaggttaacttctgctccacacgcccgtctcaacccctgc 3952  
Qy 3956 caacttctgcttgaagagactgataaaccagtgacatgatgttccactcaactatgtat 4015  
Db 3953 caacttctgcttgaagagactgataaaccagtgacatgatgttccactcaactatgtat 4012  
Qy 4016 tgaagaagactgggtgcgcgaagagacctatacagacataccgctgactgagagaagtgtac 4075  
Db 4013 tgaagaagactgggtgcgcgaagagacctatacagacataccgctgactgagagaagtgtac 4072  
Qy 4076 ctggttcaactgaacgaacagctatgtgtgtgaaggttaagagatggtgtggcgcgagt 4135  
Db 4073 ctggttcaactgaacgaacagctatgtgtgtgaaggttaagagatggtgtggcgcgagt 4132  
Qy 4136 ggtggaacgggaccccgacagatcttggtgcagacagccctgcggaaggaattcaagcgaaa 4195  
Db 4133 ggtggaacgggaccccgacagatcttggtgcagacagccctgcggaaggaattcaagcgaaa 4192  
Qy 4196 ggtcgtagctatggtccctacagcaagcttggcggtgcgcgaaggaattccataaacat 4255  
Db 4193 ggtcgtagctatggtccctacagcaagcttggcggtgcgcgaaggaattccataaacat 4252  
Qy 4256 ttatacgacacgacaggtatgtcccttgcgactgcacacgttacacggggccatctataaca 4315  
Db 4253 ttatacgacacgacaggtatgtcccttgcgactgcacacgttacacgttggccatctataaca 4312  
Qy 4316 aagggtgtgtcttaacctcagcagggagaggaataaagaacaaaggaataattcaaacct 4375  
Db 4313 aagggtgtgtcttaacctcagcagggagaggaataaagaacaaaggaataattcaaacct 4372  
Qy 4376 attgaagaccttacaatttgcacaaaaggttagctatatacactgttctgtgacatacga 4435  
Db 4373 attgaagaccttacaatttgcacaaaaggttagctatatacactgttctgtgacatacga 4432  
Qy 4436 agccaagaatctcatatctagaaggaacccagatggtcgtgacggggttgcgaagcagcagc 4495  
Db 4433 agccaagaatctcatatctagaaggaacccagatggtcgtgacggggttgcgaagcagcagc 4492  
Qy 4496 ccaaggtgttaccttgcctcaataaagacccaagcccccaagcccccaagcagcagc 4555  
Db 4493 ccaaggtgttaccttgcctcaataaagacccaagcccccaagcccccaagcagcagc 4552  
Qy 4556 gtaacacctagaagaactgacgaagataaagaataagacaggttctctgagactccgga 4615  
Db 4553 gtaacacctagaagaactgacgaagataaagaataagacaggttctctgagactccgga 4612  
Qy 4616 ggggaacctgtatacctcatatgtggaaggaatcctgtccacacaaagaaggttagata 4675

Db 4613 agggacctgtatacctcagaatgggaaggaatcctgtcccccacaaagaaggttagata 4672  
Qy 4676 tgtccaacagatatacgttcttaacccccttagagaacttaaacaccttgaaagatgtgtcag 4735  
Db 4673 tgtccaacagatatacgttcttaacccccttagagaacttaaacaccttgaaagatgtgtcag 4732  
Qy 4736 aacatcccttatactgttctgaagcttaccagaggtgtgtcagctcgtgtgtcaacaatgt 4795  
Db 4733 aacatcccttatactgttctgaagcttaccagaggtgtgtcagctcgtgtgtcaacaatgt 4792  
Qy 4796 tgttccctgacagctgtgttaactgttaactccctccagaatacctccagaagaagactaa 4855  
Db 4793 tgttccctgacagctgtgttaactgttaactccctccagaatacctccagaagaagactaa 4852  
Qy 4856 ggggaagcccccaagcgctacgttgggaagaggttaacttaactgaggttaagccccctata 4915  
Db 4853 ggggaagcccccaagcgctacgttgggaagaggttaacttaactgaggttaagccccctata 4912  
Qy 4916 cggaaacaaatatactatgttcttctgagacaccttccagatgtgtgtagagcttacc 4975  
Db 4913 cggaaacaaatatactatgttcttctgagacaccttccagatgtgtgtagagcttacc 4972  
Qy 4976 tactaagaagaagacttcaacgltgtgtgctaaagaaataactggaaggaattttccaag 5035  
Db 4973 tactaagaagaagacttcaacgltgtgtgctaaagaaataactggaaggaattttccaag 5032  
Qy 5036 attggaatacctaaggtaataagggttaagacaatggtccagcttctgttccagagtaag 5095  
Db 5033 attggaatacctaaggtaataagggttaagacaatggtccagcttctgttccagagtaag 5092  
Qy 5096 tcaaggtactgtccaaatatttgggagattgttggaaactgcatgtgtcacaacagaccca 5155  
Db 5093 tcaaggtactgtccaaatatttgggagattgttggaaactgcatgtgtcacaacagaccca 5152  
Qy 5156 aagctcagagacaggttagagagagatgaaatagaacattaaagaaccttaacaaattgac 5215  
Db 5153 aagctcagagacaggttagagagagatgaaatagaacattaaagaaccttaacaaattgac 5212  
Qy 5216 cacagaagactggttaataatgttgatgtgtccctgcgccccttgttcttttgggttga 5275  
Db 5213 cgcggaagacggtgttaataatgttgatgtgtgtccctgcgccccttgttcttttgggttga 5272  
Qy 5276 gaacacccctggaacagtttgggtgacccccctatgatatgtcttaacgggggaccccc 5335  
Db 5273 gaacacccctggaacagtttgggtgacccccctatgatatgtcttaacgggggaccccc 5332  
Qy 5336 gttggcagaatactgaccttgcacatagttgtgtgtgtgtcttccacgcttgttctc 5395  
Db 5333 atgtgtagaataatgtcttctgtacatagttgtgtgtgtgtgtcttccacgcttgttctc 5392  
Qy 5396 taagctcaagcgctcgaaggtgggtgagcgacggtggaagcagctccgggaagccta 5455  
Db 5393 taagctcaagcgctcgaaggtgggtgagcgacggtggaagcagctccgggaagccta 5452  
Qy 5456 ctc---aggaagagacttgcgaacttccacacatcgtcttccaaagtgtgagatctcaatgt 5512  
Db 5453 ctcgaagaggggagacttgcgaacttccacacatcgttccaaagtgtgagatctcaatgt 5512  
Qy 5513 tagagccacccgtgcaggaacaccccgagactcgtgtgaagggcccttaccctgtacttt 5572  
Db 5513 tagagccacccgtgcaggaacaccccgagactcgtgtgaagggcccttaccctgtacttt 5572  
Qy 5573 gaccacacaaagcggtgttaaaatgtgaaggaatccccacacttgatccatgcatccacgt 5632  
Db 5573 gaccacacaaagcggtgtgtgaaggtgaaggaatccccacacttgatccatgcatccacgt 5632  
Qy 5633 taagccggcgccacccctccgacgtcgggtgtgagagcggaagaagactgtgaatccctttaa 5692  
Db 5633 taagccggcgccacccctccgacgtcgggtgtgagagcggaagaagactgtgaatccctttaa 5692  
Qy 5693 gcttgcctccatgcgtgttccctactctcaacaaataactcccaaggcagtagtaaac 5752  
Db 5693 gcttgcctccatgcgtgttccctactctgtcaataaacctcccaaggcagtagtaaac 5752



QY	5753	gcctatagacagctcgaaccccatagaccttattcccttaoctggtatattgacc	5812
Db	5753	gcataagagacagcctcgaaccccatagaccttattcccttaoctggtatattgacc	5812
QY	5813	ctgaataagagtgatcgaataatagcgaagtggtgtcttaagagcagctgtgac	5872
Db	5813	ctgaataagagtgatcgaataatagcgaagtggtgtcttaagagcagctgtgac	5872
QY	5873	ctgaactgcatcttcgtcctcgaattgaatacccgctgtttaaagcaacaccccaac	5932
Db	5873	ctgaactgcatcttcgtcctcgaattgaatacccgctgtgtttaaagcaacaccccaac	5932
QY	5933	tagccgtagttagtggtctcattgctgcccag--sacagaagaagaatactggtg	5989
Db	5933	tagccgtagttagtggtctcattgctgcccag--sacagaagaagaatactggtg	5989
QY	5930	tcctcatgtcagcgaatttatgttgcctcgaagacacacaataatggaatacatgacg	5989
Db	5930	tcctcatgtcagcgaatttatgttgcctcgaagacacacaataatggaatacatgacg	5989
QY	5990	ggaggtctcgggaatcctctgttagagaatgagctgctcaccctccacagatgagact	6049
Db	5990	ggaggtctcgggaatcctctgttagagaatgagctgctcaccctccacagatgagact	6049
QY	6050	ggaatgcccagagattcttctgttaacaatggaactgttaacctcaactgagatatt	6049
Db	6050	ggaatgcccagagattcttctgttaacaatggaactgttaacctcaactgagatatt	6049
QY	6050	ggaatgcccagagattcttctgttaacaatggaactgttaacctcaactgagatatt	6049
Db	6050	ggaatgcccagagattcttctgttaacaatggaactgttaacctcaactgagatatt	6049
QY	6099	-ttccgcccggcgaaglacaaatgataactataaagaataagagctgctcccat	6157
Db	6110	gctctggaacaaatattactgacccgtgattagactggaagcccaatgctctctct	6169
QY	6158	cagaactagatattcgaataagatttact--gaaaggaacacaggaataattcaaa	6215
Db	6170	cagaactagatattcgaataagatttact--gaaaggaacacaggaataattcaaa	6229
QY	6216	agtgagataatgagatagcggggaataattttatattatgscggggag--cag	6271
Db	6230	aatggataatgagatagcggggaataattttatattatgscggggag--cag	6289
QY	6272	ggtcactttaaccattcgccttagagacgagacgagacacccctgttgcaattg	6331
Db	6290	ggtcactttaaccattcgccttagagacgagacgagacacccctgttgcaattg	6346
QY	6332	gaccgataaagctactgctgaacagggcccccggccctggaacacacgataactg	6391
Db	6347	gaccgataaagctactgctgaacagggcccccggccctggaacacacgataactg	6405
QY	6392	cgtgcccccaataaccctgctgagccatgaacacacgagccgctagaacagatga	6451
Db	6406	cgtgcccccaataaccctgctgagccatgaacacacgagccgctagaacagatga	6415
QY	6452	ctgagattgattcctacacacgacgctagaacacccaggtgtctctgttaagaagac	6511
Db	6416	ctgagattgattcctacacacgacgctagaacacccaggtgtctctgtgttaagaagac	6460
QY	6512	agagactccttagctcgaacagggagcttccagacacacacacacacacacacac	6571
Db	6461	caaaacttcttagcctcgaacagggagcttccagacacacacacacacacacacac	6520
QY	6572	ccacttctctgt	6621
Db	6521	ccacttctctgt	6580
QY	6632	aagaaagaanaattcaatgtgacacacacacacacacacacacacacacacacac	6691
Db	6581	gaagagggaattcaatgtgacacacacacacacacacacacacacacacacac	6640
QY	6692	ataagctac	6751
Db	6641	ataagctac	6700
QY	6752	ccac	6811
Db	6701	ccac	6760
QY	6812	tagtactgttataacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	6871
Db	6761	tagtactgttataacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	6820
QY	6872	cctcagcttcaac	6931
Db	6821	cctcagcttcaac	6880
QY	6932	actac	6991
Db	6881	actac	6940
QY	6992	gagaaac	7051
Db	6941	gagaaac	7000
QY	7052	gaacagggac	7111
Db	7001	gaacagggac	7060
QY	7112	tacatgctgac	7171
Db	7061	tacatgctgac	7120
QY	7172	agtcctgac	7231
Db	7121	agtcctgac	7180
QY	7232	tcttaagaagaagtggt	7291
Db	7181	tcttaagaagaagtggt	7240
QY	7292	actcagaagac	7351
Db	7241	actcagaagac	7300
QY	7352	gggaaagaagaagtggt	7411
Db	7301	gggaaagaagaagtggt	7360
QY	7412	ccac	7471
Db	7361	ccac	7420
QY	7472	gacctgtccttaataataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7531
Db	7421	gacctgtccttaataataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7480
QY	7532	tcatgttacttaggac	7591
Db	7481	tcatgttacttaggac	7537
QY	7592	cttcacagctccttaataataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7651
Db	7538	cttcacagctccttaataataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	7597
QY	7652	aatgcaacctaacctccagaac	7676
Db	7598	aatgcaacctaacctccagaac	7622
RESULT	5		
AA67023			
ID	AA67023	standard; DNA; 7873 BP.	
XX	AA67023;		
XX	27-MAR-2001	(first entry)	
XX	PERV env protein coding sequence SEQ ID NO: 23.		
XX	Xenotransplantation; infectious agent; vaccine; ds.		





Db 1620 aagaagataaagactaaggaagacggggttcggaacccccctcccgagccagggtaactt 1679  
QY 2206 tgaaggtggaaggggcaaccaggttgagttcctgtgtatccggagcggaacatccagtc 2265  
Db 1680 tgaaggtggaaggggcaaccaggttgagttcctgtgtatccggagcggaacatccagtc 1739  
QY 2266 tgcatacaacattaggaataactaaagaataaataatcctgtgtatccggagcggaacag 2325  
Db 1740 tactacagccattaggaataactaaagaataaataatcctgtgtatccggagcggaacag 1799  
QY 2326 aacggcagatccatggaactaccggaagacggttgacttggaggtggaacggtaacc 2385  
Db 1800 aacacagatccatggaactaccggaagacggttgacttggaggtggaacggtaacc 1859  
QY 2386 actggtttgtatccctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2445  
Db 1860 actggtttgtatccctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1919  
QY 2446 agatggagagcccaattcttcttgaacaaggaagacagagtgctgtgataaacaac 2505  
Db 1920 agatggagagcccaattcttcttgaacaaggaagacagagtgctgtgataaacaac 1979  
QY 2506 ccaatcactgtgtacccctccaattagatgataatcgactatctccccaagtaa 2565  
Db 1980 ctatccactgtgtacccctccaattagatgataatcgactatctccccaagtaa 2039  
QY 2566 agctgtacagatatacaagtcctgtgtgagacaggttccccaagcctggcgagaacgg 2625  
Db 2040 agctgtacagatatacaagtcctgtgtgagacaggttccccaagcctggcgagaacgg 2099  
QY 2626 caggagatgggtttggaacaaggaagttccccaagcgttatcaactgaaagccagtgcta 2685  
Db 2100 caggagatgggtttggaacaaggaagttccccaagcgttatcaactgaaagccagtgcta 2159  
QY 2686 caccagatcatcagtaagacagatcccttgatagaagagtcgggaaggaattggccgc 2745  
Db 2160 caccagatcatcagtaagacagatcccttgatagaagagtcgggaaggaattggccgc 2219  
QY 2746 atgttcaagaattatccaagaagcagtcagttccctgacccaactcccttggaatactc 2805  
Db 2220 atgttcaagaattatccaagaagcagtcagttccctgacccaactcccttggaatactc 2279  
QY 2806 cccgtgtaacgggttaggaagcctgggaacaaatgaattacgaacagtaacagacttgaag 2865  
Db 2280 cccgtgtaacgggttaggaagcctgggaacaaatgaattacgaacagtaacagacttgaag 2339  
QY 2866 aggtcaataaagaaggttgagacatacaaccaaggtcccgaaacccctataactcttga 2925  
Db 2340 aggtcaataaagaaggttgagacatacaaccaaggtcccgaaacccctataactcttga 2399  
QY 2926 ggcgcctcccgctggaagacgtgtaacagatctggaacttgaagaagtcctcttc 2985  
Db 2400 ggcgcctcccgctggaagacgtgtaacagatctggaacttgaagaagtcctcttc 2459  
QY 2986 ggcgtgagattacaccccaatgacacacacttctgcttcgaaatgagagatccagta 3045  
Db 2460 ggcgtgagattacaccccaatgacacacacttctgcttcgaaatgagagatccagta 2519  
QY 3046 cgggaagaacccgggagactcaactcctggagccgagcggcccaaggttcaagaactcccgga 3105  
Db 2520 cgggaagaacccgggagactcaactcctggagccgagcggcccaaggttcaagaactcccgga 2579  
QY 3106 ccatctttgagagagccctacagggagcctggaacttgaagatcccaacccctag 3165  
Db 2580 ccatctttgagagagccctacagggagcctggaacttgaagatcccaacccctag 2639  
QY 3166 tgaacctcctcagtaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3225  
Db 2640 tgaacctcctcagtaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2699  
QY 3226 tagaagtagaagagcactactgctgtgaattgtctgagactaggtacagagcctctgcta 3285  
Db 2700 tagaagtagaagagcactactgctgtgaattgtctgagactaggtacagagcctctgcta 2759  
QY 3286 agaagggccagatttgcaagagagatgaataacttgggtgagtgagtgagtgagtgag 3345  
Db 2760 agaagggccagatttgcaagagagagatgaataacttgggtgagtgagtgagtgagtgag 2819  
QY 3346 agcagatgctgagcagagagcagaggaagaaactgtatccagaatcccgcccaacag 3405  
Db 2820 agcagatgctgagcagagagcagaggaagaaactgtatccagaatcccgcccaacag 2879  
QY 3406 ccaaaaagtgagagaggttttgggagacgctggatcttgcagactgtggatcccggggt 3465  
Db 2880 ccaaaaagtgagagaggttttgggagacgctggatcttgcagactgtggatcccggggt 2939  
QY 3466 ttgcagcttaagcagcccaactcaccgcttaacaaagaagaaggggaggtctcctgg 3525  
Db 2940 ttgcagcttaagcagcccaactcaccgcttaacaaagaagaaggggaggtctcctgg 2999  
QY 3526 ctccctgagacagaagagcatttgatctataaaaaggccctgtgagcagcctgtcc 3585  
Db 3000 ctccctgagacagaagagcatttgatctataaaaaggccctgtgagcagcctgtcc 3059  
QY 3586 tggcctccctgagcagtaacttaaacctttaccccttatgtgagtgagtgagtgagtgag 3645  
Db 3060 tggcctccctgagcagtaacttaaacctttaccccttatgtgagtgagtgagtgagtgag 3119  
QY 3646 cccgagagattttaaacccaacccctagagccaatggagagacgtgtgctacatgca 3705  
Db 3120 cccgagagattttaaacccaacccctagagccaatggagagacgtgtgctacatgca 3179  
QY 3706 agaagcttatcctgtagccagtggttggcccgatgtgtgagagctatccagctgtgg 3765  
Db 3180 agaagcttatcctgtagccagtggttggcccgatgtgtgagagctatccagctgtgg 3239  
QY 3766 ccaactgtgtcaagagcgtctgacaataatgaacttggagagagataaactgtatagccc 3825  
Db 3240 ccaactgtgtcaagagcgtctgacaataatgaacttggagagagataaactgtatagccc 3299  
QY 3826 cccatgcatltggagaacacatcgttccgcaagcccccagacagatggaatgacaacggcgca 3885  
Db 3300 cccatgcatltggagaacacatcgttccgcaagcccccagacagatggaatgacaacggcgca 3359  
QY 3886 tgaaccaatatacaagcgtcttctacagaagaaggtcaacttgcctccacagcgctc 3945  
Db 3360 tgaaccaatatacaagcgtcttctacagaagaaggtcaacttgcctccacagcgctc 3419  
QY 3946 tcaaccttccactcttccgctggaagagacgtatgaacacagtgactgaatgacatc 4005  
Db 3420 tcaaccttccactcttccgctggaagagacgtatgaacacagtgactgaatgacatc 3479  
QY 4006 aactatgtattgagagacgtgggttcgcaaggaaccttaacagataccgctgactggag 4065  
Db 3480 aactatgtattgagagacgtgggttcgcaaggaaccttaacagataccgctgactggag 3539  
QY 4066 aagtgtaacgtgttcaactgagcaggaagcagctatgtatggaaggttaagagatggcgt 4125  
Db 3540 aagtgtaacgtgttcaactgagcaggaagcagctatgtatggaaggttaagagatggcgt 3599  
QY 4126 gggcggcagtggtggaacgggacccgacgacgtggtggaagcgtggaaggaact 4185  
Db 3600 gggcggcagtggtggaacgggacccgacgacgtggtggaagcgtggaaggaact 3659  
QY 4186 cagcgcaaaagcctgagcctatgaccccaagccttggcgtggaaggaact 4245  
Db 3660 cagcgcaaaagcctgagcctatgaccccaagccttggcgtggaaggaact 3719  
QY 4246 ccaataacattatacagcagcaggtatgtccttggcagtcacacgtacacggggcca 4305  
Db 3720 ccaataacattatacagcagcaggtatgtccttggcagtcacacgtacacggggcca 3779  
QY 4306 tctataaagaaggggttggcttaacctcagcagggggaagaaataagaacaaaggaagaa 4365  
Db 3780 tctataaagaaggggttggcttaacctcagcagggggaagaaataagaacaaaggaagaa 3839

QY 4366 ttctaagcctattagaagccttacatttgcacaaagcgtagctattatacactgtcctg 4425  
 Db 3840 ttctaagcctattagaagccttacatttgcacaaagcgtagctattatacactgtcctg 3899  
 QY 4426 gacatcagaagcacaagatctctatatctatagaggaacacagatggtgaccgggttgcca 4485  
 Db 3900 gacatcagaagcacaagatctctatatctatagaggaacacagatggtgaccgggttgcca 3959  
 QY 4486 agcagcagccagcgtgttaaccttctctctatataatagaacacgcccacagaac 4545  
 Db 3960 agcagcagccagcgtgttaaccttctctctatataatagaacacgcccacagaac 4019  
 QY 4546 ccagacagcagtlacaccttagaagactggcaagagataaaaaagatagaccagtctctg 4605  
 Db 4020 ccgagcagcagtlacaccttagaagactggcaagagataaaaaagatagaccagtctctg 4079  
 QY 4606 agactccgagggagcctgctatacctctatctggaaggaatactgccccacaagaag 4665  
 Db 4080 agactccgagggagcctgctatacctctatctggaaggaatactgccccacaagaag 4139  
 QY 4666 ggttagaataatgtccaaacagatcacctctctaacccacttagaactaaacacctgcagc 4725  
 Db 4140 ggttagaataatgtccaaacagatcacctctctaacccacttagaactaaacacctgcagc 4199  
 QY 4726 agttgtcagaacatcccccttatctgttctgaggtctacagagtggtgactcggtg 4785  
 Db 4200 agttgtcagaacatcccccttatctgttctgaggtctacagagtggtgactcggtg 4259  
 QY 4786 tcaaacatgtgtccctgcccagctggttaatgctaaactctccagaaatacctccagaa 4845  
 Db 4260 tcaaacatgtgtccctgcccagctggttaatgctaaactctccagaaatacctccagaa 4285  
 QY 4846 agagactaagggaagccacccagcgctcactgggaagtggacttcaactgaggtaaagc 4905  
 Db 4286 -----ggtaaaagc 4293  
 QY 4906 cggttaaatlacggaacaaatactatgtgttttttagacacccctttcagatgggtgag 4965  
 Db 4294 cggttaaatlacggaacaaatactatgtgttttttagacacccctttcagatgggtgag 4353  
 QY 4966 aggtctatactactaagaagagacttaacccgtggtgctaaagaaaaactctgga-ggaa 5024  
 Db 4354 aggtctatactactaagaagagacttaacccgtggtgctaaagaaaaactctggaagaa 4413  
 QY 5025 attttccaagatttgaataccttaaggttaaggtgctagacaaatggtccagcttccgt 5084  
 Db 4414 attttccaagatttgaataccttaaggttaaggtgctagacaaatggtccagcttccgt 4473  
 QY 5085 gccaggttaagtcagggaactggccaaagatttggggattggtggaactgcatgtgca 5144  
 Db 4474 gccaggttaagtcagggaactggccaaagatttggggattggtggaactgcatgtgca 4533  
 QY 5145 tacagacccaaagctcaggacaggttagagaggaataatagaacattaaagacccctt 5204  
 Db 4534 tacagacccaaagctcaggacaggttagagaggaataatagaacattaaagacccctt 4593  
 QY 5205 accaaattgcccacagagactggcatttaattgattggtggtctcctccttctgtctt 5264  
 Db 4594 actaaattgcccacagagactggcatttaattgattggtggtctcctccttctgtctt 4653  
 QY 5265 tttagggtgaggaacccctgacaggtttgggctgaccccttatgaaattgctctacggg 5324  
 Db 4654 tttagggtgaggaacccctgacaggtttgggctgaccccttatgaaattgctctacggg 4713  
 QY 5325 ggaccccccctgttgccagaaattgctttgacatagtgctgattgctgttctccag 5384  
 Db 4714 ggaccccccctgttgccagaaattgctttgacatagtgctgattgctgttctccag 4773  
 QY 5385 cctttgtctctaggtcgaagcgctcagtggtggtgagcgagcggtgggaagcgctc 5444  
 Db 4774 cctttgtctctaggtcgaagcgctcagtggtggtgagcgagcggtgggaagcgctc 4833  
 QY 5445 cgggagggcctactc---aggagagacttgcaaggttccaatcgtctccaaagtggagat 5501

Db 4834 cgggagggcctactcaggagggagagacttgcagatccacatcgtttccaaagtgggagat 4893  
 QY 5502 tcaagtctatgttagacgcccacccgtgcagagaaacctcgagactcgttggagggagacattat 5561  
 Db 4894 tcaagtctatgttagacgcccacccgtgcagagaaacctcgagactcgttggagggacccctat 4953  
 QY 5562 ctctgaacttttgacacacacacacgctgtgaaagtgcgaagaaatcccccactggtatccat 5621  
 Db 4954 cactgaacttttgacacacacacacgctgtgaaagtgcgaagaaatcccccactggtatccat 5013  
 QY 5622 gcatccacgtttaagcggcgccacccctcccgactcggggtggagagccgaaaaagactgag 5681  
 Db 5014 gcatccacgtttaagcggcgccacccctcccgactcggggtggagagccgaaaaagactgaa 5073  
 QY 5682 aatcccttaagcttgcctccatccctggttctctactactaataaactccccagggc 5741  
 Db 5074 aatcccttaagcttgcctccatccctggttctctactacttcaataaactccccaggt 5133  
 QY 5742 cagtagtaaacgcctta tagacagctcgaaaccccatagacactttatcccttaactggtc 5801  
 Db 5134 taatgttaaacgccttggagacggcgaactccataaaaccccttatctcacttggtt 5193  
 QY 5802 gattattgacctgatacaggggttoactgttaaatagcaactcgaaggtgtgtcctcagagg 5861  
 Db 5194 acttactgactccgtacaggtattatattaaacagcactcaagggaggtccctctggg 5253  
 QY 5862 cactgtgtggtgactgaacttctctcctccgattgattaaacccctgtta-----a 5915  
 Db 5254 gactgtgtggtgactgaacttctctcctcagtcagtcagtcagtcagtcagtcagtcagtc 5313  
 QY 5916 aagcacactcccaacactgactcgttattgggttctcttattgtgcccagg---cacaga 5972  
 Db 5314 ggcacaccccccgatgaactcgttactcgttactcgttactcgttactcgttactcgtt 5373  
 QY 5973 gaaagagaataactgt 6032  
 Db 5374 taatgaagaataattgtgaaatcctcagatttcttttgcagcaatggagctgcgttaac 5433  
 QY 6033 ctccaacgtatgagactgaaatggcgtatctctcctcagacgggtgtaaatctctctt 6092  
 Db 5434 ttctaagtgtgggaattgggaatggcgtatctctcagcaacagagatgaattactcttt 5493  
 QY 6093 tgtcaattcc-----g 6103  
 Db 5494 tgttaacaactcctaccagttataatcaatttaattatggcctggagatggaaagattg 5553  
 QY 6104 gccgggcaagtacaaaatgatgaactataataagataagagctgtccccatcagact 6163  
 Db 5554 gcaacagcgggtacaaaagatg-tacgaaataagcaataaagctgtcattcgttagacc 5612  
 QY 6164 tagattatctaaagataagtttctactg--aaaggaaaaacagaaaaatttcaaaagtga 6221  
 Db 5613 tagattatctaaagataagtttctactgaaaaagaaaaacagaaaaatttcaaaagtga 5672  
 QY 6222 taaatggtatgagctggggaatagttttttat-----tatatggoggggagcaggggtcca 6277  
 Db 5673 taaatggtatgcttggggaatagtttacttggaggtcttggggaagaaagagatctg 5732  
 QY 6278 ctttaacattcgtcttaggatagagacggggagacagacccctgtgcaatggagcccg 6337  
 Db 5733 ttctgactattcgtctcagaa tagaaactcagatggaaacctcgttctctatagaccacaa 5792  
 QY 6338 ataaagtagctgctgaacagggggcccgccctggagccacgcataacttgcoggtgc 6397  
 Db 5793 ataaaggttggccgaacaaagacactccaatccaaga----- 5829  
 QY 6398 cccaattaaactcgtcgtcgccctgacataacacagcccgcttagcgaacagtagtaccactgagat 6457  
 Db 5830 -----acagagggccatctcctaaacctctgattacaatacaacctgagat 5876  
 QY 6458 tgattcctacaaacgcctagaaaaactccccagggttctcctgttaagacagagacagac 6517

Db 5877 cagtcgccactga-----gcttaacatcactattataaaacagggcgaaac 5921  
Qy 6518 tcttaagttctccagggaggttcccaagccatcaactccaccagacctgatgcactt 6577  
Db 5922 tttttaacctccagggaggttcccaagctcttaactccacgactccagagctacct 5981  
Qy 6578 ctctctgttggttctctatctccagggctccttattatagagggttggtctaaagaa 6637  
Db 5982 ctctctgttggttctctatctccagggctccttattatagagggttggtctaaagaa 6641  
Qy 6638 gaaatccaatgtgacaaagagcatagaaatcaatgtacatgggggtcccgaaataagc 6697  
Db 6042 gaaatccaatgtgacaaagagcaatagagccaatgtacatgggggtcccgaaataagc 6101  
Qy 6698 ttacctactgaagtctccgggaaggagcatgcataggaagagctcccccaccacc 6757  
Db 6102 ttacctactgaagtctccgggaaggagcatgcataggaagagctcccccaccacc 6161  
Qy 6758 aacacctttgctatagtagtctgtttatagagcgggctccgaaataatagtagtac 6817  
Db 6162 aacacctttgctatagtagtctgtttatagagcgggctccgaaataatagtagtac 6221  
Qy 6818 ctggtttaacacaggttggtgcatgcaatactggtttaacccctgtgttccaacctcag 6877  
Db 6222 ctggtttaacacaggttggtgcatgcaatactggtttaacccctgtgttccaacctcag 6281  
Qy 6878 tcttaacccaatccaaagatttctgtctatggtcccaatgctcccccagagtgactacc 6937  
Db 6282 tcttaacccaatccaaagatttctgtctatggtcccaatgctcccccagagtgactacc 6341  
Qy 6938 atcttgaggaagtggtctctgataatgactatgactatgactatgactatgactatg 6997  
Db 6342 atccgaaaaagagctccttgatgaatgactatgactatgactatgactatgactatg 6401  
Qy 6998 ccgtatcccttaacctagctaatgctgattaggaagcgcgttgggtaggaacag 7057  
Db 6402 ccatactctgacactagctaatgctgatttggagtggttgcaggcgttgggaacag 6461  
Qy 7058 ggacgctgcccctgatcacagagccacagcagctagagaagagcttggtagctacatg 7117  
Db 6462 gaaagctgcccctaatcacagagccacagcagctggaagagcttagtaacctacatc 6521  
Qy 7118 cggccatgacagaagatctccagccttaaaaggagctgtgtagcaacctagaagagctccc 7177  
Db 6522 gaattgaacgggaagatctccagccttagaaaaatctgtcagtaacctgaggaatccc 6581  
Qy 7178 tgactctcttctggaagtggtctacagaaaccgagaggtttagatctgtcttcttaa 7237  
Db 6582 taacctctctatggaagtggtctacagaaacagaaagggtttagatctgttcttaa 6641  
Qy 7238 gagaaggttggttctgtgagccttaaaagaagaatgttctctatgtagatcaactcag 7297  
Db 6642 aagaagaggattatgtgtagccttgaaggaggaatctgtttttatgtggaattcag 6701  
Qy 7298 gagccatcagagactccatgaacaagcttgaaaaaa----- 7334  
Db 6702 gggccatcagagactccatgaacaagcttagaagaagactggagagcgcccggtct 6761  
Qy 7335 ----- 7334  
Db 6762 cgaacacccagacaggttctgttcttaataaagaactgtcgaagtaaccgctgagct 6821  
Qy 7335 -----gttagagagcgctcgaagggaagaaagagaggtgacccaggggt 7375  
Db 6822 aaagccagcttagagaaggttggagagcgtcgaagggaagaaactactcaagagt 6881  
Qy 7376 ggtttgaagaggttccaaaggtctccttgatgagcaccacctgtcttctgtctgacgg 7435  
Db 6882 ggtttgagaggttccaaaggtctccttggttggtggtggtggtggtggtggtggtggt 6941  
Qy 7436 gggcccttagtagtctcctctcttacttacttacttacttacttacttacttacttactt 7495  
Db 6942 gacccttaatagtagtctcctctcttacttacttacttacttacttacttacttactt 7001

Qy 7496 ttgctttgttagagaacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 7555  
Db 7002 ttgctttcattagagaacgaataagtgagtgagtgagtgagtgagtgagtgagtgagtgag 7061  
Qy 7556 aaggccttctgagccaggagaaactgacctctagccttccaggttcttaagtagaact 7615  
Db 7062 aagcccgcttagc---agagaagctggcgctagctctaccaggttcttaagtagaact 7118  
Qy 7616 attaaacaagaagaagtgagggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaa 7675  
Db 7119 attaaacaagaagaagtgagggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaa 7178  
Qy 7676 ccagggaagttaataaaagctctaaatgcccgcgaatgaaatgaaatgaaatgaaatgaaatgaa 7735  
Db 7179 ccagggaagttaataaaagctctaaatgcccgcgaatgaaatgaaatgaaatgaaatgaaatgaa 7232  
Qy 7736 aatagtagaaggttcacaccttctctatgttccagggcctgtctatcctggcctaagtag 7795  
Db 7233 --ataggtaaaagatcatcacttttctgtttttagggttcttctgtctgtacaaaa 7290  
Qy 7796 ataacaggaatgagttga----- 7814  
Db 7291 ctttgggaagggaagaaacagggccctctgagtgatgtgctctctatgcttgaacctcttga 7350  
Qy 7815 -----ctaactgcttctggttctgtaaaactgactg----- 7848  
Db 7351 aactgctctaaactggttctgttctgttctgttctgttctgttctgttctgttctgttctgt 7410  
Qy 7849 -----gcaccatagaagaattgattacacattgacacgacctgtgacc 7891  
Db 7411 gaagtcaattgcttaacgagcccccagtaagatcggtgtaccacaaaatgttgaacaca 7470  
Qy 7892 tatctcaactgcaatgtctactctctccagagagccacgagatcgagacctccggagc 7951  
Db 7471 tatctgttgacaacatgtctcccccaccccaacacatgcgaaatgtgtaacctctaaaa 7530  
Qy 7952 tattttaaaatgattgtccacggagcgcggtctctcgatattttaaaatgattgttoca 8011  
Db 7531 caatttaaaatgattgtccacggagcgcggtctctcgagttttaaaatgactggt--- 7587  
Qy 8012 tggagcgcggtctctcgatattttaaaatgattgttctgacgcacagggcttctgtg 8071  
Db 7588 -----ttgtgattttgaaatgattgttctgaaagcggttctgttctgtgta 7635  
Qy 8072 aaccccataaaagctgtcccgactcccgactcggggcgagctctaccctcggtg 8131  
Db 7636 aaccccataaaagctgtcccgactcccgactcggggcgagctctaccctcggtg 7695  
Qy 8132 tgcagactgtggcccccagcgcttggaataaaaaatcctctgtctgtttgcatcaaaa 8191  
Db 7696 tgcagactgtggcccccagcgcttggaataaaaaatcctctgtctgtttgcatcaaga 7755

## RESULT 6

AAC67019  
ID AAC67019 standard; DNA; 7362 BP.

XX AAC67019;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 19.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.





QY	3956	cactcttgcctgaagagactgatgaaccagtgactcatgatgtgccaatacactattgat	4015
Db	4616	cactcttgcctgaagagactgatgaaccagtgactcatgatgtgccaatacactattgat	4675
QY	4016	tgaagagactgggttcgcaaggaaccttacagacataccgtgactgagaagaagtgtcaac	4075
Db	4676	tgaagagactgggttcgcaaggaaccttacagacataccgtgactgagaagaagtgtcaac	4735
QY	4076	ctgttcaactgaaggagcagctatgtgtgaaaggttaagaggatggctggggcgccaggt	4135
Db	4736	ctgttcaactgaaggagcagctatgtgtgaaaggttaagaggatggctggggcgccaggt	4795
QY	4136	ggttgacggaaccgcacagatctggccagcagctgcgggaagaaccttcagcgcaaaa	4195
Db	4796	ggttgacggaaccgcacagatctggccagcagctgcgggaagaaccttcagcgcaaaa	4855
QY	4196	ggctgagctcatggcctcacgaagctttgcgctggccgaagggaatccataaaacat	4255
Db	4856	ggctgagctcatggcctcacgaagctttgcgctggccgaagggaatccataaaacat	4915
QY	4256	ttatacgacagcaggtatgcctttgcgactgcacacgtacacggggccatctataaaca	4315
Db	4916	ttatacgacagcaggtatgcctttgcgactgcacacgtacacggggccatctataaaca	4975
QY	4316	aagggggtgtcttacctcagcaggagggaataaagaacaaagagagaattcttaagcct	4375
Db	4976	aagggggtgtcttacctcagcaggagggaataaagaacaaagagagaattcttaagcct	5035
QY	4376	attagaagccttacatttgcgaagagctagctattatacacttgtcctggacatcagaa	4435
Db	5036	attagaagccttacatttgcgaagagctagctattatacacttgtcctggacatcagaa	5095
QY	4436	agccaaagatctcatctatagagggaacacagatggctgacccgggttcgcaagcaggcagc	4495
Db	5096	agccaaagatctcatctatagagggaacacacagatggctgacccgggttcgcaagcaggcagc	5155
QY	4496	ccagggtgttaacctctgcctataatagaaccccaaaagcccaagcccaagccagacgaca	4555
Db	5156	ccagggtgttaacctctgcctataatagaaccccaaaagcccaagcccaagccagacgaca	5215
QY	4556	gtacaccttagaagactggcaagagataaaaagatagacagcttctctgagactccgga	4615
Db	5216	gtacaccttagaagactggcaagagataaaaagatagacagcttctctgagactccgga	5275
QY	4616	ggggacctgctatacctcatatgggaagaaatcctgcccaaaaaagggtagaata	4675
Db	5276	agggacctgctatacctcatatgggaagaaatcctgcccaaaaaagggtagaata	5335
QY	4676	tgtcccaacagatatacgtctaacccactaggaaactaaacactcacagctggtctcag	4735
Db	5336	tgtcccaacagatatacgtctaacccactaggaaactaaacactcacagctggtctcag	5395
QY	4736	aacatcccttatactgttctgaggtaccaggagtggtgactcgggtgtcaaacatttg	4795
Db	5396	aacatcccttatactgttctgaggtaccaggagtggtgactcgggtgtcaaacatttg	5455
QY	4796	tgtgccctgccagctggtaattgctaatccttccagataactccaggaagaagagactaag	4855
Db	5456	tgtgccctgccagctggtaattgctaatccttccagataactccaggaagaagagactaag	5515
QY	4856	gggaagccaccagcgctcactcgggaagtggacttcaactgaggttaaagccggctaaata	4915
Db	5516	gggaagccaccagcgctcactcgggaagtggacttcaactgaggttaaagccggctaaata	5575
QY	4916	cggaaacaaatatctattgtgtttttgtagacaccttttcaggatgggtgaggtctatccc	4975
Db	5576	cggaaacaaatatctattgtgtttttgtagacaccttttcaggatgggtgaggtctatccc	5635
QY	4976	tactaagaagaagactcaaccggtgtgtgctaagaagaataactggagaagaattttccaaag	5035
Db	5636	tactaagaagaagactcaaccggtgtgtgctaagaagaataactggagaagaattttccaaag	5695

QY	5036	atttggatacctaaggtataatagggtcagacaataaggctccagcttctcgttgccaggtaag	5095
Db	5696	atttggatacctaaggtataatagggtcagacaataaggctccagcttctcgttgccaggtaag	5755
QY	5096	tcagggaactggccaagatatgggggattgattggaaactgcaattgtgtacacagaccoca	5155
Db	5756	tcagggaactggccaagatatgggggattgattggaaactgcaattgtgtacacagaccoca	5815
QY	5156	aagctcaggacagatgagagagatgaatagaaccattaaagagaccctaccaaattgac	5215
Db	5816	aagctcaggacagatgagagagatgaatagaaccattaaagagaccctaccaaattgac	5875
QY	5216	cacagagactggcattaatgatggatgggtctcctgcctttgtgcttttttagggtag	5275
Db	5876	cgcggagactggcgttaattgatggatgggtctcctgcctttgtgcttttttagggtag	5935
QY	5276	gaacacccctggacagattgggctgacccctatgaattgctctacggggagaccccccc	5335
Db	5936	gaacacccctggacagattgggctgacccctatgaattgctctacggggagaccccccc	5995
QY	5336	gttggcgaataatgccttttgacatatgtctgatgtgtgtgttccagccttgttctc	5395
Db	5996	atttgtagaataitgcttcogtaacatgctgtgacgtgctgttccagccttgttctc	6055
QY	5396	taggtctcaagcgtcgtcgtggttgaggcagcgagcgtggagcagctccggaggacct	5455
Db	6056	taggtctcaagcacttgatgggttgaggcacaacgagcgtggaggcaactccggaggacct	6115
QY	5456	ctc--aggaggagacttgcaagtccacatcgcttccaaagtggagattcagtcctatgt	5512
Db	6116	ctcaggaggagagacttgagatcccatcctcctggttccaaagtggagattcagtcctatgt	6175
QY	5513	tagcgcacacgttgacgaaacctcgagactcgttggaaggagaccttatctcgtactttt	5572
Db	6176	tagcgcacacgttgcaaggaaacctcgagactcgttggaaggagaccttatctcgtactttt	6235
QY	5573	gaccacaccacgcgttgaaagtgcgaaggaaatccccacctggatcgcattccaccgt	5632
Db	6236	gaccacaccacgcgttgaaagtgcgaaggaaatccccacctggatcgcattccaccgt	6295
QY	5633	taagccggcgcacacctcccgacttcgggttggaagaccgaaagaagactgagaatcccttaa	5692
Db	6296	taagccggcgcacacctcccgacttcgggttggaagaccgaaagaagactgagaatcccttaa	6355
QY	5693	gcttcgctccatcgcctggttcttaactcttaacataactcccgagccagtagtaaac	5752
Db	6356	gcttcgctccatcgcctggttcttaactctgtaataactctcctcaggttaaaggttaaac	6415
QY	5753	gccttatagacagctgcgaacccccatagacctttatcccttacctgcgtgattatgacc	5812
Db	6416	gccttggagacgccgaactccataaaccttatctcctcaactggttacttactgact	6475
QY	5813	ctgatacgggtgtcactgttaatatagcacttcaggtgtgtctcctagaggcaccttggcgc	5872
Db	6476	ccggtagaggtatataataacagcactcaaggggagggctcccttgggacctgttggcgc	6535
QY	5873	ctgaactgattctcctccgatgattaaaccgcgttta-----aaagcaacacct	5926
Db	6536	ctgaattatgtctgcctctcagtaactcccttggctcctcaatgacagggccacacccc	6595
QY	5927	ccaacctagtcctgattatgggttcttatgtgcgccagg---cacagagaaagagaaat	5983
Db	6596	ccgatgactccgtgcttacgggttttacgtttgcccagacccccccaataatgaagaat	6655
QY	5984	actgtgggggttctggggaatccttctgttaggagatggagctgcgtcacctcccaacgat	6043
Db	6656	attgtggaactcctcaggattcttttgcagagcaatggagctggttaactctctaaatgagt	6715
QY	6044	gaacitggaaatggcgaatctctcctccaggaccgggttaaaattctcctttgt-----	6095
Db	6716	ggaaatggaaatggccagctctcctcagcaacagacagtagtactctttttgttaacaatc	6775
QY	6096	-----caattccgcgcgcggcaag	6114



Db	6776	ctaccagttataatcaatttaattatggccatgggaagtggaaagattgggcaacacg	ggg	6833
QY	6115	tacaaaatgatgaactatataaagataaagctgctcccatcagacttagattatcta	6174	
Db	6836	tacaaaagatg-tacgaaataagcaataagctgtcatctgctagacctagattacta	6894	
QY	6175	aagataagtttcactg--aaagaaaaacagaaaataattcaaaagtggaataatggtatg	6232	
Db	6895	aaataaagtttcactgaaaagaaaacagaataattcaaaagtggaataatggtatg	6954	
QY	6233	agctgggaatagttttttat----tatatgcggggagcaggtgccactttaaccatt	6288	
Db	6955	tcttgggaataagtgtactatggagctctgggagaaagaaagatctgtctgactatt	7014	
QY	6289	cgccttaggataagaacggggacagaacccctctggcaatgggacccgataaagtactg	6348	
Db	7015	cgcctcagaataagaactcagatgaagacctcggttctatagaccacaataaggttctg	7074	
QY	6349	gtgtaacaggggcccccggccctggagcacgcataactgcccgtgcccaataaaccc	6408	
Db	7075	gccgaacaaggccctccaatccaagaaca-----	7103	
QY	6409	tgcctggcctgtacataaacacagccgcctagcaaacagtcaccactgattgattccacc	6468	
Db	7104	----gaggccattctctaaacccctctgattacaatacaacctctggatccaccact	7158	
QY	6469	aacagcctagaaactccccagggttctcctgttaagacagacagagactcttcagcttc	6528	
Db	7159	ga-----gcctaacatcacattaaaacagggcgaaactttttaacctc	7203	
QY	6529	atccaggagctttccaagccatacaactcaacgacctgatgccacttctttgttgg	6588	
Db	7204	atccaggagcttttcagctcttaactccaagcactccagaggtacctctctgttgg	7263	
QY	6589	ctttgctatctcagccctctctattatgaggggatggtctaaagaaagaaattcaat	6648	
Db	7264	ctttgctagtctgggccaccttactatgatgaggaatggctgagaggagaaattcaat	7323	
QY	6649	gtgacaaagacatagaaaatcaatgtacatgggggtcc	6687	
Db	7324	gtgacaaaggaacatagagacaaatgtacatggggatcc	7362	

## RESULT

RESULT	
AAAF77726	
ID	AAAF77726 standard; DNA; 7333 Bp.
XX	
AC	
XX	AAAF77726;
XX	
DT	23-MAY-2001 (first entry)
XX	
DE	Defective retroviral genome isolated from PK-15 cell line.
XX	
XX	
KW	Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
KK	ss.
XX	
OS	Unidentified.
XX	
PN	US6190861-B1.
XX	
PD	20-FEB-2001.
XX	
PF	13-DEC-1996; 96US-0766528.
XX	
PR	14-DEC-1995; 95US-0572645.
XX	
PA	(GEHO ) GEN HOSPITAL CORP.
XX	
PI	Fishman JA;
XX	
DR	WPI; 2001-256211/26.
DR	P-PSDB; AAF73282, AAF73283, AAF73284.
DR	



QY	721	cattcgatggttgatggccatcagagggaacotttaattctgaaattatctcgtgctgta	780
Db	731		790
		cattcgatggttgatggccatcagagggaacottlaattctgaaattatctcgtgctgta	
QY	781	aggcaatcatttttcagactggaccggctctcatctctgatacaggagccctatacttta	840
Db	791		850
		aggcaatcatttttcagactggaccggctctcatctctgatacaggagccctatacttta	
QY	841	cgfgccaagatttggcagaagatctctccggcaatgggttaaacctaggctaataaaccaa	900
Db	851		910
		cgfgccaagatttggcagaagatctctccggcaatgggttaaacctaggctaataaaccaa	
QY	901	gaaagcagatcccgaatctctggctcttggagagaaaaaaacaaactcggccgaaaaag	960
Db	911		970
		gaaagcagatcccgaatctctggctcttggagagaaaaaaacaaactcggccgaaaaag	
QY	961	tcgagccctctctcgatactaacccgagatcggaggagccgcgacttggccggaaacc	1020
Db	971		1029
		tcgagccctc-tcctcgatactaacccgagatcggaggagccgcgacttggccggaaacc	
QY	1021	caacctgttcccccccccttatccagacagggtgctgtgagggga-cctctggccct	1079
Db	1030		1089
		caacctgttcccccccccttatccagacagggtgctgtgaggggaacctctggccct	
QY	1080	ccctggagctccggtggtgagaggacctgctccgggaactcggagcccgagagagccacc	1139
Db	1090		1149
		ccctggagctccggtggtgagaggacctgctccgggaactcggagcccgagagagccacc	
QY	1140	ccggagcgacagacagagatcgatattaccctgcgacactatgccctcccatgccca	1199
Db	1150		1209
		ccggagcgacagacagagatcgatattaccctgcgacactatgccctcccatgccca	
QY	1200	ggggggccaattgcagccctccagttatggccctttctctgcagatctctataattgg	1259
Db	1210		1269
		ggggggccaattgcagccctccagttatggccctttctctgcagatctctataattgg	
QY	1260	aaactaaacaatccccttctcggaggatccccaacgctcagccgggttgggtgagttcc	1319
Db	1270		1329
		aaactaaacaatccccttctcggaggatccccaacgctcagccgggttgggtgagttcc	
QY	1320	cttatgtctctcaccagctacttgggatgattgcaacagctgctgcagacaactcttc	1379
Db	1330		1389
		cttatgtctctcaccagctacttgggatgattgcaacagctgctgcagacaactcttc	
QY	1380	acaaacggagcgagagagaattctctgttagagctagaaaaatgttctggggcgac	1439
Db	1390		1449
		acaaacggagcgagagagaattctctgttagagctagaaaaatgttctggggcgac	
QY	1440	ggcgaccacacagcttgcaaaatgagatgaatggatttcccttgactcggccggt	1499
Db	1450		1509
		ggcgaccacacagcttgcaaaatgagatgaatggatttcccttgactcggccggt	
QY	1500	tgggactcaaacagcgctgaagtgaggagagcttgaatactatccaggtcttggtg	1559
Db	1510		1569
		tgggactcaaacagcgctgaagtgaggagagcttgaatactatccaggtcttggtg	
QY	1560	gcgggtctccggggcgctcaagacggcccaactaatltggctaaggtaagagaggtgatg	1619
Db	1570		1629
		gcgggtctccggggcgctcaagacggcccaactaatltggctaaggtaagagaggtgatg	
QY	1620	cagggaccgaacgaacctccctcgatttcttgagaggtctatggaggtcttcaggcgg	1679
Db	1630		1689
		cagggaccgaacgaacctccctcgatttcttgagaggtctatggaggtcttcaggcgg	
QY	1680	ttcaccccttttgactactcactcagagggcccaagaagcctcagttggccctggcttcatt	1739
Db	1690		1749
		ttcaccccttttgactactcactcagagggcccaagaagcctcagttggcccttcatt	
QY	1740	gggcagctcggtctctggatatacgaagaagaaactcagagactggaaggtttacaggggct	1799
Db	1750		1809
		gggcagctcggtctctggatatacgaagaagaaactcagagactggaaggtttacaggggct	

QY	1800	gagttacgtgactagtgagagagcgagagaagtgtattacagagagggagagacagagag	1859
DB	1810	gagttcacgtgactcagtgagagagcgagagaagtggtattacagaagggagacagagag	1869
QY	1860	gagaggaacagagaaaaagagagagagagaagggagagaaagcgtgatagacgg	1919
DB	1870	gagaaaggcaagagaaaaagagagagagaagaaagggagagaagacgtgatagacgg	1929
QY	1920	caagagaagaatttgactaagatcttggccgcagtggttgaagggaagacagcagcgag	1979
DB	1930	caagagaagaatttgactaagatcttggccgcagtggttgaagggaagacagcagcgag	1989
QY	1980	agagagagagattttagaaaaatttaggtcaaggccctagacagtcagggaaacctgggcaat	2039
DB	1990	agagagagagattttagaaaaatttaggtcagggccctagacagtcagggaaacctgggcaat	2049
QY	2040	aggagcccaactcgacaagaccagtgtcggtattgtaaaagaaaaagacactgggcaag	2099
DB	2050	aggagcccaactcgacaagaccagtgtcggtattgtaaaagaaaaagacactgggcaag	2109
QY	2100	aactgccccaagaaggaaaaaagacccaagtgctcagtcctagaagaagataaagat	2159
DB	2110	aactgccccaagaaggaaaaaagacccaagtgctcagtcctagaagaagataaagat	2169
QY	2160	tagggagagcgggttcggaccacctcccgagcccgagtcgaactttgaagtgtagggg	2219
DB	2170	tagggagagcgggttcggaccacctcccgagcccgagtcgaactttgaagtgtagggg	2229
QY	2220	caaccagttgagttctctggttgataccggagcggagacattcaagtctcgtacaaccatta	2279
DB	2230	caaccagttgagttctctggttgataccggagcggagacattcaagtctcgtacaaccatta	2289
QY	2280	ggaaactaaagaaaaaaaactcctgggttgatgggtgcacacggcgcaacgcagtatcca	2339
DB	2290	ggaaactaaagaaaaaaaactcctgggttgatgggtgcacacggcgcaacgcagtatcca	2349
QY	2340	tggactaccggaagaaacggttgacttgggagtgggacgggttaaccacctgtttctggct	2399
DB	2350	tggactaccggaagaaacggttgacttgggagtgggacgggttaaccacctgtttctggct	2409
QY	2400	atccctgagtgcccagtaacctctctaggtagagacttactgtaccaagataggagctcaa	2459
DB	2410	atccctgagtgcccagtaacctctctaggtagagacttactgtaccaagataggagctcaa	2469
QY	2460	attctttgaacaagaagaaacagcagaagtgctgtgtaatacaaaaacctacactgtgtg	2519
DB	2470	attctttgaacaagaagaaacagcagaagtgctgtgtaatacaaaaacctacactgtgtg	2529
QY	2520	acctccaaattagatgatgaatacgaactatattctcccaagtaaacgctgtacaagat	2579
DB	2530	acctccaaattagatgatgaatacgaactatattctcccaagtaaacgctgtacaagat	2589
QY	2580	atacagtccttggttgagacagtttcccgaagcctgggcagaaaaacgcagggatgggttg	2639
DB	2590	atacagtccttggttgagacagtttcccgaagcctgggcagaaaaacgcagggatgggttg	2649
QY	2640	gcaagcaagttcccacacaggttattcaactgaaggccagtgctcacacagatcaagtc	2699
DB	2650	gcaagcaagttcccacacaggttattcaactgaaggccagtgctcacacagatcaagtc	2709
QY	2700	agacagtaccctctgagttagagaggtcgcagaaggaatttggcgcgatgttcaagaatta	2759
DB	2710	agacagtaccctctgagttagagaggtcgcagaaggaatttggcgcgatgttcaagaatta	2769
QY	2760	atccaaacagggcactcgtgtcctgtccaatccccttggaatactcccctgctaccgggtt	2819
DB	2770	atccaaacagggcactcgtgtcctgtccaatccccttggaatactcccctgctaccgggtt	2829
QY	2820	aggagcctgggaccaatgattatcgaccagtcacaggacttgagagaggttcaataaagg	2879
DB	2830	aggagcctgggaccaatgattatcgaccagtcacaggacttgagagaggttcaataaagg	2889
QY	2880	gtgcaggacatacacccaaacggctcccgaaaccttataacctcttgagcgccctccgcct	2939

Db 2890 |||||gtcaggacatacccccacggtcccgaaaccttataaacctcttgagcgcctcccgct 2949  
QY 2940 gaacgaaactggtacacagttatgacttaaaagatgcttctctctgctgagattacac 2999  
Db 2950 |||||gtcaggaaactggtacacagttatgacttaaaagatgcttctctctgctgagattacac 3009  
QY 3000 cccactagccaaaccttttgccttcgaattgagagatccaggtacgagggaacccggg 3059  
Db 3010 cccactagccaaaccttttgccttcgaattgagagatccaggtacgagggaacccggg 3069  
QY 3060 cagctcaactggaccgcgactccccaagggttcaagaactccccgacctctttgacgaa 3119  
Db 3070 cagctcaactggaccgcgactccccaagggttcaagaactccccgacctctttgacgaa 3129  
QY 3120 gccctacacaggacctgccaacttcaggtacgatacccaacctcaggtgacctctccag 3179  
Db 3130 gccctacacaggacctgccaacttcaggtacgatacccaacctcaggtgacctctccag 3189  
QY 3180 tacgtggatgacctgtctctgcgaggagcccaaacacaggactgcttagaagggtacgaag 3239  
Db 3190 tacgtggatgacctgtctctgcgaggagcccaaacacaggactgcttagaagggtacgaag 3249  
QY 3240 gaactactgctggaattgctgacctaggtacagagcctctgcttaagaaggccagatt 3299  
Db 3250 gaactactgctggaattgctgacctaggtacagagcctctgcttaagaaggccagatt 3309  
QY 3300 tqcaggagaggttaacatacttgggttacagtttgcggggcgagcgatgctgacg 3359  
Db 3310 tqcaggagaggttaacatacttgggttacagtttgcggggcgagcgatgctgacg 3369  
QY 3360 gaggcacggaagaaacttgtagtcagataccggcccaacacacagcccaacaagtga 3419  
Db 3370 gaggcacggaagaaacttgtagtcagataccggcccaacacacagcccaacaagtga 3429  
QY 3420 gagtttttgggaagctggatttgcagactgtgatcccggggttgcgaccttagca 3479  
Db 3430 gagtttttgggaagctggatttgcagactgtgatcccggggttgcgaccttagca 3489  
QY 3480 gccccactctaccgcttaacaaagaaaggggattctctcgtggctcgtgagcaccag 3539  
Db 3490 gccccactctaccgcttaacaaagaaaggggattctctcgtggctcgtgagcaccag 3516  
QY 3540 aaggcatttgatgctatacaaaaggccctgtgagcgcaactgtctgtgcccctccctgac 3599  
Db 3517 ----- 3516  
QY 3600 gtaactaaacctttacccctttatgtgtagagcgttaaggagtagcccgagggtttta 3659  
Db 3517 ----- 3516  
QY 3660 acccaaaccttaggaccatggaggagacctgttgctacctgtcaagaaagcttgatcct 3719  
Db 3517 ----- 3516  
QY 3720 gtacgcagtggttggccccgtatgtctgaaggtatcgagctgtgcccactactggtcaag 3779  
Db 3517 ----- 3516  
QY 3780 gacgctgacaaatgactttgggacagataataactgtaataagccccccatgcatggag 3839  
Db 3517 ----- 3516  
QY 3840 aacatcgttcggcagccccccagacggtggtgacccaacgcccgcgcatgacccactatcaa 3899  
Db 3517 ----- 3516  
QY 3900 agcctgtctctcaacagagagggtcactttcgtctccaccagccgctctcaacccctgcaact 3959  
Db 3517 ----- 3516  
QY 3960 cttctgctgaagagactgatgaaccagtgactcatgattgattgattgag 4019

Db 3517 ----- 3516  
QY 4020 gagctggggtccgcgaaggaccttatacagatacccgctgactggagaagtgtctaacctgg 4079  
Db 3517 ----- 3516  
QY 4080 ttcactgacggaagcagcgtatgtggtggaaggttaagagatggtcgtggcgagtcggtg 4139  
Db 3517 ----- 3516  
QY 4140 gacgggaccgcgcagatcttgccgacgagcctgccgaaggaacttcagcgcaaaagcct 4199  
Db 3517 ----- 3516  
QY 4200 gagctcatggccctcaccgcaagctttgcggtgcccgaagggaatccataaaacatttat 4259  
Db 3517 ----- 3516  
QY 4260 acggacagcaggtatgcctttgcgactgcacgtaacggtgcacgggcccctctataaaacaaagg 4319  
Db 3517 -----aaagg 3521  
QY 4320 ggttgccttacctcagcagcagggaaataaaacaaagagaaattcttaagccttata 4379  
Db 3522 ggttgccttacctcagcagcaggggaaataaaacaaagagaaattcttaagccttata 3581  
QY 4380 gaagccttacatttgcacaaaggctagctattatatacactgtcctggacatcagaagcc 4439  
Db 3582 gaagccttacatttgcacaaaggctagctattatatacactgtcctggacatcagaagcc 3641  
QY 4440 aaagatctcatctagaggaacccagatggtcagccgggttcccaagcagcagccag 4499  
Db 3642 aaagatctcatctagaggaacccagatggtcagccgggttcccaagcagcagccag 3701  
QY 4500 gctgttaaccttctgcctataatagaaaccccaagcccaagcccaagacacagctac 4559  
Db 3702 gctgttaaccttctgcctataatagaaaccccaagcccaagacacagctac 3761  
QY 4560 acctagaagactggcgaagagataaaagacagcttctctgagactccggagggg 4619  
Db 3762 acctagaagactggcgaagagataaaagacagcttctctgagactccggagggg 3821  
QY 4620 acctgtataacctcaataggaaggaatacctcccccaaaaggggttagaatgtc 4679  
Db 3822 acctgtataacctcaataggaaggaatacctcccccaaaaggggttagaatgtc 3881  
QY 4680 caacagatacatcgtctaacccaccttaggaactaaacacctgcagcagttggtcagaaca 4739  
Db 3882 caacagatacatcgtctaacccaccttaggaactaaacacctgcagcagttggtcagaaca 3941  
QY 4740 tcccccttatcgtttctgaggtaccagggagtgctgactcgtgtgctaaacatttgttg 4799  
Db 3942 tcccccttatcgtttctgaggtaccagggagtgctgactcgtgtgctaaacatttgttg 4001  
QY 4800 ccttgccagctggttaagtataccttccagaatacctccagaaagagactaaagggga 4859  
Db 4002 ccttgccagctggttaagtataccttccagaatacctccagaaagagactaaagggga 4061  
QY 4860 agccaccagcgcctcactggagtggaacttcaactgaggtaaagccgctaaatcggga 4919  
Db 4062 agccaccagcgcctcactggagtggaacttcaactgaggtaaagccgctaaatcggga 4121  
QY 4920 acaaaatctatttgggtttttagacaccttttcaggatgggtgagacttctact 4979  
Db 4122 acaaaatctatttgggtttttagacaccttttcaggatgggtgagacttctact 4181  
QY 4980 aagaaagagacttcaaccgtggtggttaagaaataactggagaaatttttccaaattt 5039  
Db 4182 aaaaaagagacttcaaccgtggtggttaagaaataactggagaaatttttccaaattt 4241  
QY 5040 ggaatacctaaggttaaggttcagacaaatggtccagcttctggtccaggttaagtcag 5099  
Db 4242 ggaatacctaaggttaaggttcagacaaatggtccagcttctggtccaggttaagtcag 4301

[illegible]



CC cells, infectivity titration and susceptibility testing can be  
CC performed. ultimately animal donors without intact porcine retroviral  
CC sequences or a lower copy number of viral elements could be selected.  
XX  
SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

Query Match 66.0%; Score 5416.6; DB 18; Length 7393;  
Best Local Similarity 82.4%; Pred. NO. 0;  
Matches 6854; Conservative 0; Mismatches 404; Indels 1065; Gaps 13;

QY 1 gtggttaoactgtggccccagcgcttggaataaaaaatcctctgtgttgcate 60  
DB 11 gtggttaoactgtggccccagcgcttggaataaaaaatcctctgtgttgcate 70  
QY 61 aagacgccttcgtgaagtattgggtgtcgcccttcacgagccgagaggggat 120  
DB 71 aagacgccttcgtgaagtattgggtgtcgcccttcacgagccgagaggggat 130  
QY 121 tgtctttactggcctttcatttggctgtggcggaatacctcgacacccctta 180  
DB 131 tgtctttactggcctttcatttggctgtggcggaatacctcgacacccctta 190  
QY 181 caccagagaacacttgaggtaaaggatccctttggaacgtgtgtgtgcgcc 240  
DB 191 caccagagaacacttgaggtaaaggatccctttggaacatgtgtgtgcgcc 250  
QY 241 ggcgtctctgttcgagtcgtcttttcgtgatgcgcctttcggttcgagctgcct 300  
DB 251 ggcgtctctgttcgagtcgtcttttcgtgatgcgcctttcggttcgagctgcct 310  
QY 301 ctgagaccgttaagactggagactgtatcagcagacgtgtcagagatcacaggctg 360  
DB 311 ctgagaccgttaagactggagactgtatcagcagacgtgtcagagatcacaggctg 370  
QY 361 ceacccctggggagcccccggagggtggggagagccaggaagcctgtgtgtctact 420  
DB 371 ceacccctggggagcccccggagggtggggagagccaggaagcctgtgtgtctact 430  
QY 421 gtcggttcagagaccagatctgtgttgaagcgaagccttcccctccgcccgcgcg 480  
DB 431 gtcggttcagagaccagatctgtgttgaagcgaagccttcccctccgcccgcgcg 490  
QY 481 actcttttgcctgtgttgaagacgcgagcgggtcgctgtgtcggatctgtgtgttt 540  
DB 491 actcttttgcctgtgttgaagacgcgagcgggtcgctgtgtcggatctgtgtgttt 550  
QY 541 ctgtctcgtgtcttctgtctgtcgtctgtctacagttttaatatgggacagacag 600  
DB 551 ctgtctcgtgtcttctgtctgtcgtctgtctacagttttaatatgggacagacag 610  
QY 601 tgactacccctttagttgactctcgaccattgactgaagttagatccagggctcata 660  
DB 611 tgactacccctttagttgactctcgaccattgactgaagttagatccagggctcata 670  
QY 661 attgtcagttcaggttaagaaggacattggcagacttctgtcctctgaatggccaa 720  
DB 671 attgtcagttcaggttaagaaggacattggcagacttctgtcctctgaatggccaa 730  
QY 721 cattcagatttgatggccatcagaggggacattttaaattcgtaaaattccttggttta 780  
DB 731 cattcagatttgatggccatcagaggggacattttaaattcgtaaaattccttggttta 790  
QY 781 aggcacatcattttcagactgagcccgctctcactcctgatcagagccctatacctta 840  
DB 791 aggcacatcattttcagactgagcccgctctcactcctgatcagagccctatacctta 850  
QY 841 cgtggcaagatttggcagaagatcctccgcatgggtttaaaccatggcttaataaccaa 900  
DB 851 cgtggcaagatttggcagaagatcctccgcatgggtttaaaccatggcttaataaccaa 910  
QY 901 gaaagccaggtcccccgaatccttgctcttggagagaaaaaacacactcggccgaaaaag 960  
DB 911 gaaagccaggtcccccgaatccttgctcttggagagaaaaaacacactcggccgaaaaag 970

DB 911 gaaagccaggtcccccgaatccttgctcttggagagaaaaaacacactcggccgaaaaag 970  
QY 961 tcgagccctcttctctctgtatctaccccgagatcgagggagccgcaacttggccgaaacc 1020  
DB 971 tcgagccctc-tctctgtatctaccccgagatcgagggagccgcaacttggccgaaacc 1029  
QY 1021 caactctgttcccccaccccccttatccagcacagggtgtgtgagggga-cctctgcacct 1079  
DB 1030 caactctgttcccccaccccccttatccagcacagggtgtgtgagggacacctctgcacct 1089  
QY 1080 cctggagctccggtgtgtgagggagcctgtccgggactcggagccgagagggcgccacc 1139  
DB 1090 cctggagctccggtgtgtgagggagcctgtccgggactcggagccgagagggcgccacc 1149  
QY 1140 ccggagcgagacagacagatcgcgatatatccgctgcgacactatggccctcccatgcc 1199  
DB 1150 ccggagcgagacagacagatcgcgatatatccgctgcgacactatggccctcccatgcc 1209  
QY 1200 gggggccaattgcagccctccagtattggcccttttctctgcagatctctataattgg 1259  
DB 1210 gggggccaattgcagccctccagtattggcccttttctctgcagatctctataattgg 1269  
QY 1260 aaaaataaccatccccctttctcgagatccccaacgctcaccgggtgtgtgagtgcc 1319  
DB 1270 aaaaataaccatccccctttctcgagatccccaacgctcaccgggtgtgtgagtgcc 1329  
QY 1320 ctatgtctctccacgacctacttgggatgttgcacagctgtcgcagacactcttc 1379  
DB 1330 ctatgtctctccacgacctacttgggatgttgcacagctgtcgcagacactcttc 1389  
QY 1380 acaaccgagagcagacagaaattctgttagaggtgagaaaaatgttctggggcgac 1439  
DB 1390 acaaccgagagcagacagaaattctgttagaggtgagaaaaatgttctggggcgac 1449  
QY 1440 gggcgaccacgcgagtgcaaaatgagattgacatgggatttcccttgactcgcgcgggt 1499  
DB 1450 gggcgaccacgcgagtgcaaaatgagattgacatgggatttcccttgactcgcgcgggt 1509  
QY 1500 tgggactacaacagcgctgaaggtagggagagcttgaatactatcgcaggtctggtg 1559  
DB 1510 tgggactacaacagcgctgaaggtagggagagcttgaatactatcgcaggtctggtg 1569  
QY 1560 gcgggtctccggggcgccctcaagcgcccactaatttggctaaaggttaagagagtgatg 1619  
DB 1570 gcgggtctccggggcgccctcaagcgcccactaatttggctaaaggttaagagagtgatg 1629  
QY 1620 caggagccgaacgaacctcccctcggtatttcttgagaggtctatggagccttcagcg 1679  
DB 1630 caggagccgaacgaacctcccctcggtatttcttgagaggtctatggagccttcagcg 1689  
QY 1680 ttcaccccttttgatcctacatcagagggccagaaagcctcagtgcccttgccctcatt 1739  
DB 1690 ttcaccccttttgatcctacatcagagggccagaaagcctcagtgcccttgccctcatt 1749  
QY 1740 gggcagtcgctctggatatcagaaagaaacttcagagactggaggtttacagagagct 1799  
DB 1750 gggcagtcgctctggatatcagaaagaaacttcagagactggaggtttacagagagct 1809  
QY 1800 gagttacgtgtatctagtggagggcagagaggtgtattacagaaagggagacagagag 1859  
DB 1810 gagttacgtgtatctagtggagggcagagaggtgtattacagaaagggagacagagag 1869  
QY 1860 gagaaggaacagagaaaaagagagagagagagagagagagagagagagagagag 1919  
DB 1870 gagaaggaacagagaaaaagagagagagagagagagagagagagagagagagag 1929  
QY 1920 caagagaagaatttgcataagatcttggccagctgttgaaggaagagacagcaggag 1979  
DB 1930 caagagaagaatttgcataagatcttggccagctgttgaaggaagagacagcaggag 1989  
QY 1980 agagagagagatttttagaaaaattaggttcagggcctgacagctcagggaaacctgggcaat 2039  
DB 1990 agagagagagatttttagaaaaattaggttcagggcctgacagctcagggaaacctgggcaat 2049

QY 2040 agaacccactgcacaaagaccagtgctgtattgttaagaaagacactggcgcaag 2099  
|||||  
Db 2050 agaacccactgcacaaagaccagtgctgtattgttaagaaagacactggcgcaag 2109  
|||||  
QY 2100 aactgccccagaagggaaacaaagaccaggtctctagtagaagaataaagat 2159  
|||||  
Db 2110 aactgccccagaagggaaacaaagaccaggtctctagtagaagaataaagat 2169  
|||||  
QY 2160 tagggagacgggttcggaccctcccagaccaggttaactttgaaggtgagggg 2219  
|||||  
Db 2170 tagggagacgggttcggaccctcccagaccaggttaactttgaaggtgagggg 2229  
|||||  
QY 2220 caaccagttgagttcctgttgataccggagcgagcattcagtgctctacaaccatta 2279  
|||||  
Db 2230 caaccagttgagttcctgttgataccggagcgagcattcagtgctctacaaccatta 2289  
|||||  
QY 2280 ggaacataaagaaataatctcgtgtggtgacacagggcaacggcagatatacca 2339  
|||||  
Db 2290 ggaacataaagaaataatctcgtgtggtgacacagggcaacggcagatatacca 2349  
|||||  
QY 2340 tgactaccgaagacgttgactggagtgagtgagcaggttaacccactggttctggtc 2399  
|||||  
Db 2350 tgactaccgaagacgttgactggagtgagtgagcaggttaacccactggttctggtc 2409  
|||||  
QY 2400 atccctgagtgccagtagccctcttagtagtagacttactgacaaagatgggagctcaa 2459  
|||||  
Db 2410 atccctgagtgccagtagccctcttagtagtagacttactgacaaagatgggagctcaa 2469  
|||||  
QY 2460 attcttttgaaaggaagaccagagtgctgtgataaacaacccactcaactgtgttg 2519  
|||||  
Db 2470 attcttttgaaaggaagaccagagtgctgtgataaacaacccactcaactgtgttg 2529  
|||||  
QY 2520 accctccaatagatgataatcgactatattctcccagtaagacccctgatacaagat 2579  
|||||  
Db 2530 accctccaatagatgataatcgactatattctcccagtaagacccctgatacaagat 2589  
|||||  
QY 2580 atacagtcctggttgagcagttctcccaagcctgggcagaaacccgagggatgggttg 2639  
|||||  
Db 2590 atacagtcctggttgagcagttctcccaagcctgggcagaaacccgagggatgggttg 2649  
|||||  
QY 2640 gaaagcaagttcccccaggttattcaactgaagccagtgctacacagtagtcagtc 2699  
|||||  
Db 2650 gaaagcaagttcccccaggttattcaactgaagccagtgctacacagtagtcagtc 2709  
|||||  
QY 2700 agacgtacccttgagtagagagctcgagaagaaattggccgcagtgttcaagatta 2759  
|||||  
Db 2710 agacgtacccttgagtagagagctcgagaagaaattggccgcagtgttcaagatta 2769  
|||||  
QY 2760 atccaaagggatcctagttcctgtccaatcccttggaataactccctgctaccgggt 2819  
|||||  
Db 2770 atccaaagggatcctagttcctgtccaatcccttggaataactccctgctaccgggt 2829  
|||||  
QY 2820 aggaagcctgggaccaaagattatcgaccagtagacaggtgagaggtcaataaaagg 2879  
|||||  
Db 2830 aggaagcctgggaccaaagattatcgaccagtagacaggtgagaggtcaataaaagg 2889  
|||||  
QY 2880 gtgcaggacatacacccacgggtcccgaaacctataacctttgagcgccctcccgcct 2939  
|||||  
Db 2890 gtgcaggacatacacccacgggtcccgaaacctataacctttgagcgccctcccgcct 2949  
|||||  
QY 2940 gaacgaactgttacacagtagttgacttaaaagatgctcttctgctgagattacac 2999  
|||||  
Db 2950 gaacgaactgtgtacacagtagttgacttaaaagatgctcttctgctgagattacac 3009  
|||||  
QY 3000 cccactagccaaccatttttgccttcgaatggagagatccaggttacggggaagaccggg 3059  
|||||  
Db 3010 cccactagccaaccatttttgccttcgaatggagagatccaggttacggggaagaccggg 3069  
|||||  
QY 3060 cagctcaactggaccgcactgccccaaaggggttcaagaactcccagccatotttgacaa 3119  
|||||  
Db 3070 cagctcaactggaccgcactgccccaaaggggttcaagaactcccagccatotttgacaa 3129  
|||||

QY 3120 gcctacacagggacactggccaaacttcaggatccaaacacccctcagtgacccctctccag 3179  
|||||  
Db 3130 gcctacacagggacactggccaaacttcaggatccaaacacccctcagtgacccctctccag 3189  
|||||  
QY 3180 tacgtgagtagcctgtcttcggcgagaccacaaacagactgcttaagagtagcaag 3239  
|||||  
Db 3190 tacgtgagtagcctgtcttcggcgagaccacaaacagactgcttaagagtagcaag 3249  
|||||  
QY 3240 gcactactgctgaaattgtctgacctaggctacagagcctctgctaaagggccagatt 3299  
|||||  
Db 3250 gcactactgctgaaattgtctgacctaggctacagagcctctgctaaagggccagatt 3309  
|||||  
QY 3300 tgcagagagagtagtaacatacttggggtagacttttcggggcgagcgatggctgacg 3359  
|||||  
Db 3310 tgcagagagagtagtaacatacttggggtagacttttcggggcgagcgatggctgacg 3369  
|||||  
QY 3360 gaggcgagaaagaaacttagtccagataccggcccccaacacagcccaaaagtgaga 3419  
|||||  
Db 3370 gaggcgagaaagaaacttagtccagataccggcccccaacacagcccaaaagtgaga 3429  
|||||  
QY 3420 gagtgtttgggagacagctgattttgcagacttgtagtcccggttttcgaccttagca 3479  
|||||  
Db 3430 gagtgtttgggagacagctgattttgcagacttgtagtcccggttttcgaccttagca 3489  
|||||  
QY 3480 gccccactctacccgctaaccaaaagaaaggggagattctctctggtcctgagcaccag 3539  
|||||  
Db 3490 gccccactctacccgctaaccaaaagaa----- 3516  
QY 3540 aaggcatttgatgctatcaaaaaggccctgctgagcgacactgctctggccctccctgac 3599  
Db 3517 ----- 3516  
QY 3600 gtaactaaacctttaccctttatgttgatgagcgtaaaggagtagcccgaggagttaa 3659  
Db 3517 ----- 3516  
QY 3660 acccaaaccttagaccatagagagacactgttgctacactgtcaaaagacttgatcct 3719  
Db 3517 ----- 3516  
QY 3720 tagccagtggttgcccgctatgtctgaaggtatcgagctgtgcccatactggtcaag 3779  
Db 3517 ----- 3516  
QY 3780 gacgtgacaaattgactttggagagataataactgtaagtagcccccactgcatggag 3839  
Db 3517 ----- 3516  
QY 3840 aacatgcttcggcagccccacagacagatggatgaccaacgcccgcagatgacccactatcaa 3899  
Db 3517 ----- 3516  
QY 3900 agcctgctctcacagagaggtgtaactttogtccaccagcgcctctcaacctgccaact 3959  
Db 3517 ----- 3516  
QY 3960 cttctgctgaagagactgatgaacacagtagactcatgattgcatcaactattgattgag 4019  
Db 3517 ----- 3516  
QY 4020 gagactgggtccgcgaaggaccttacagacataccgctgactggagaagtgttaacctgg 4079  
Db 3517 ----- 3516  
QY 4080 ttcactgacggaagcagctatgtgttggaaggtaagagtagtctgggagtgagtggtg 4139  
Db 3517 ----- 3516  
QY 4140 gacgggacccgcagcagatctggccagcagcctgccggaaggaacttcagcgcaaaagcct 4199  
Db 3517 ----- 3516  
QY 4200 gagctcagtgccctcacgcaagcttttggcgtggcggaaggaaatocataaaacattat 4259

Db 3517 ----- 3516  
QY 4260 acggacagcaggtatgctcttgcgactgcacacgtacacgggcccacatctataaaacaaagg 4319  
Db 3517 -----aaagg 3521  
QY 4320 ggggtgttacctcagcagggaggaaataaagaaacaaagagaaattctaagcctatta 4379  
Db 3522 ggggtgttacctcagcagggaggaaataaagaaacaaagagaaattctaagcctatta 3581  
QY 4380 gaagccttacattgccccaaaaaggctagctattatcacactgtcctggacatcagaagacc 4439  
Db 3582 gaagccttacattgccccaaaaaggctagctattatcacactgtcctggacatcagaagacc 3641  
QY 4440 aaagatctcatatcagagggaaccagatggctgaccgggttgccaagcaggacccacag 4499  
Db 3642 aaagatctcatatcagagggaaccagatggctgaccgggttgccaagcaggacccacag 3701  
QY 4500 gctgttaaccttctgctataataagaaacgccccaaagccccagacccagacagctac 3761  
Db 3702 gctgttaaccttctgctataataagaaacgccccaaagccccagacccagacagctac 3761  
QY 4560 accctagaagactgcagagagataaaaaagatagaccgggttctctgagactccggagggg 4619  
Db 3762 accctagaagactgcagagagataaaaaagatagaccgggttctctgagactccggagggg 3821  
QY 4620 acctgctatacctcatatgggaaggaaatcctgcccccaagaaagaggttagaataatgc 4679  
Db 3822 acctgctatacctcatatgggaaggaaatcctgcccccaagaaagaggttagaataatgc 3881  
QY 4680 caacagatatacgtctaacccacctagaaactaaacacctgcagcagttggtcagaaca 4739  
Db 3882 caacagatatacgtctaacccacctagaaactaaacacctgcagcagttggtcagaaca 3941  
QY 4740 tccccctatcatgttctgaggtaccagagtggtggtgactcgtggtggtcaaacattgttg 4799  
Db 3942 tccccctatcatgttctgaggtaccagagtggtggtgactcgtggtggtcaaacattgttg 4801  
QY 4800 cccctccagctggttaactgctaaactctccagaatacctccagaaagagactaaggga 4859  
Db 4002 cccctccagctggttaactgctaaactctccagaatacctccagaaagagactaaggga 4061  
QY 4860 agccacccagcgtcactgggaagtgcacttcaactgaggttaagcgggtcaaaatacga 4919  
Db 4062 agccacccagcgtcactgggaagtgcacttcaactgaggttaagcgggtcaaaatacga 4121  
QY 4920 acaaaatctattgtttttgtagacaccttttcaggatgggttagaggcttactact 4979  
Db 4122 acaaaatctattgtttttgtagacaccttttcaggatgggttagaggcttactact 4181  
QY 4980 aagaaagagactcaaccgtggtggctaagaaataactggaggaatttttccagattt 5039  
Db 4182 aaaaaagagactcaaccgtggtggctaagaaataactggaggaatttttccagattt 4241  
QY 5040 ggaatacctaaggttaattaggttcagacaattggtccagctttcgttccaggtgaagtca 5099  
Db 4242 ggaatacctaaggttaattaggttcagacaattggtccagctttcgttccaggtgaagtca 4301  
QY 5100 ggaactggccaagatatgtgggatttgattggaaactgattgtgatacagaccccaagc 5159  
Db 4302 ggaactggccaagatatgtgggatttgattggaaactgattgtgatacagaccccaagc 4361  
QY 5160 tcaggacaggttagagagatgaataagacattaaagaccccttaccataatgaaccaca 5219  
Db 4362 tcaggacaggttagagagatgaataagacattaaagaccccttaccataatgaaccaca 4421  
QY 5220 gagactggcattaatgattgagtggtctctcgtcccttctgcttttttaggttaggaac 5279  
Db 4422 gagactggcattaatgattgagtggtctctcgtcccttctgcttttttaggttaggaac 4481  
QY 5280 accctggacagtttggtgacccctatgaattgctctacaggggacccccccgttg 5339  
Db 5280 accctggacagtttggtgacccctatgaattgctctacaggggacccccccgttg 5339

Db 4482 accctggacagtttggttgacccccctataaaattgctctacaggggacccccccgttg 4541  
QY 5340 gcagaaattgcttcttgacatagtgctgctgctgtgttttccagcctttgttctctag 5399  
Db 4542 gcagaaattgcttcttgacatagtgctgctgctgtgttttccagcctttgttctctag 4601  
QY 5400 ctcaaggcgtcgcagtggttgagcagcagcgtggaagcagctcccgaggcctactca 5459  
Db 4602 ctcaaggcgtcgcagtggttgagcagcagcgtggaagcagctcccgaggcctactca 4661  
QY 5460 ggaggagacttgcaagttccacatcgcttccaagttggagattcgactctatgttagacgc 5519  
Db 4662 ggaggagacttgcaagttccacatcgcttccaagttggagattcgactctatgttagacgc 4721  
QY 5520 caccgtgcagaaaccttgagagactcgttggaaggacacttatctctgacttttgaccaca 5579  
Db 4722 caccgtgcagaaacctcgagactcgttggaaggacacttatctctgacttttgaccaca 4781  
QY 5580 ccaacggctgtgaaagtcgaaggaaatccccacactggatcgatccacgttaagccg 5639  
Db 4782 ccaacggctgtgaaagtcgaaggaaatccccacactggatcgatccacgttaagccg 4803  
QY 5640 gcgccacctcccgactcgggtggagagccgaaaaagactgagaaatcccccttaagcttcgc 5699  
Db 4804 -----gaatcccttaagcttcgc 4822  
QY 5700 ctccatcgctggttctctactctactaaacaataactccccagccagtagtaaaacgcttat 5759  
Db 4823 ctccatcgctggttctctactctactaaacaataactccccagcttaaatggtaaacgcttgc 4882  
QY 5760 agcagctgcgaacccccatagaccttatcccttacctggtgattattgacctgatac 5819  
Db 4883 ggacagcccgaaactccccataaaccttatctctcaccctggttacttactgacctggatc 4942  
QY 5820 ggggtgactgtaaatagcactcgaggtgtgctcctagagcagcctggtggtgacctgaact 5879  
Db 4943 aggttaataatataacagcactcaaggaggctcccttgggacctggtggtgacctgaatt 5002  
QY 5880 gcaatttgcctccgattgattaccccgctgtta -----aaagcacacactccccacct 5933  
Db 5003 atatgtcgtcctcgactcagtaactcctggtctcaatgaccagggcacaccccccgatgt 5062  
QY 5934 agtcctgatttgggttctctattgctgcgcagg ---cacagagaaagagaaatctgttg 5990  
Db 5063 actcgtgcttaagggttttctacgtttgcgcagaccccccaaaataatgaagaaatctgttg 5122  
QY 5991 ggggttcgggaatcctctgttaggagatggagctgcgtcacctccacagatggagactg 6050  
Db 5123 aaatcctcaggatttcttttgcagcaatggagctgcataaacttctaagtggggaattg 5182  
QY 6051 gaaatggcgatctctccagacccgggttaaaattctcctcttgcataatcc ----- 6102  
Db 5183 gaaatggcgatctctccagacccgggttaaaattctcctcttgcataatccaccag 5242  
QY 6103 -----ggcccgggcaggtacacaaa 6121  
Db 5243 ttataatcaatttaattatggccatggagatggaaagattggcaacagcgggtacacaaa 5302  
QY 6122 tgatgaactatataaagataagagctgctccccctcagacttagattatctaaagataa 6181  
Db 5303 agatg-tacgaataagcaataagctgctcattctgtagacctagattacttaaaaaataa 5361  
QY 6182 gtttctcgtg-aaagaaaaacagaaaaatttcaaaagtggaataatggtatgagctggg 6239  
Db 5362 gtttctcgtgaaaaagaaaaacaaaaatttcaaaagtggttaattatctctggg 5421  
QY 6240 gaatggtttttat----tatatggcgggggagcaggttccactttaacattcgcctta 6295  
Db 5422 gaatggttactatgaggtcctcggagaaaaagaaagaaatctgttctgactattcgcctca 5481  
QY 6296 ggaatagacggggagacacccctctggtgaatgggaccccgataaagactggtggaac 6355  
Db 5482 gaatagaaactcagatggacacctcgggttgctatagggacccaataaagggtttggccgaac 5541







PD 30-NOV-2000.  
 XX 24-MAY-2000; 2000WO-US14296.  
 XX 24-MAY-1999; 99US-0135631.  
 XX (MAYO-) MAYO MEDICAL VENTURES.  
 XX Federspiel MJ;  
 PT WPI; 2001-032041/04.  
 DR  
 XX  
 XX Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 XX Claim 16; Page 107-109; 144pp; English.  
 PS  
 XX  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

Query Match  
 Best Local Similarity 61.2%; Score 5025.8; DB 22; Length 6076;  
 Matches 5234; Conservative 0; Mismatches 222; Indels 15; Gaps 6;

QY 1 gtgggtgacgactgtggcccgagcgcttggaataaaatccctctgtgtgttgcac 60  
 Db  
 601 gfggtgacgactgtggcccgagcgcttggaataaaatccctctgtgttgcac 660  
 QY 61 aagacgctctcgtgagtgatt-tgggggtgcgctcttcgagccggagcgaggga 119  
 Db  
 661 aagacgctctcgtgagtgattggaagggagtcgcttttcgagcctggagg----- 713  
 QY 120 ttgtctttactgccccttattgttgcgttggccgggaaatccctcgaccaccctt 179  
 Db  
 714 --ttcttttgcgtcttatttgggggtcgtccgggtatctgtgcggccaccctta 771  
 QY 180 acacccgagacccgacttggaggt-aaagggtatccctttggaacgtgtgtg--tgtgc 236  
 Db  
 772 acacccgagacccgacttggaggtaaagagatcctcttttaacgtgtatgtatgacc 831  
 QY 237 ggcgggctctgtctgt 296  
 Db  
 832 ggcgggctctgt 891  
 QY 297 tccctcagacgttaagacttgagactgtgatcagcagacgtgtcgttagggatcacag 356  
 Db  
 892 tccctcagacgttaagacttgagactgtgatcagcagacgtgtcgttagggatcacag 951  
 QY 357 gctgcacccctggggagcggccggaggttggggagagcccgagcgtgtgtgtgtccc 416  
 Db  
 952 gctgcacccctggggagcggccggaggttggggagagcccgagcgtgtgtgtgtccc 1011  
 QY 417 tactcgtcgtcagagacccgagttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 476  
 Db  
 1012 tctcgtcgtcagagacccgagttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1071  
 QY 477 tccgactcttctgt 536  
 Db  
 1072 tccgactcttctgt 1131  
 QY 537 gtttctgtctgt 596  
 Db  
 1132 gtttctgt 1191  
 QY 597 acagtgactaccccccttagttgactctgacccattggactgaagttagatccagggt 656

Db 1192 acggtgacgacccctcttagttgactctgacccattggactgaagttaaatccagggt 1251  
 QY 657 cataattgtcagttcaggttaagaaggacacttggcagacatttctgtgcctctgaatgg 716  
 Db 1252 cataattgtcagttcaggttaagaaggacacttggcagacatttctgtgcctctgaatgg 1311  
 QY 717 ccaacattcgtattgttgatggccatcagaggggaccccttaattctgaaattatctcgtgct 776  
 Db 1312 cagacattcgtattgttgatggccatcagaggggaccccttaattctgagattatctcgtgct 1371  
 QY 777 gtttaaggcaatcttttgcagctggaccggcctctcctcctcctcctcctcctcctcctcct 836  
 Db 1372 gtttaaggcaatcttttgcagctggaccggcctcctcctcctcctcctcctcctcctcctcct 1431  
 QY 837 cttacggtgcaagatttggcagaagatcctcgcctcctcctcctcctcctcctcctcctcctcct 896  
 Db 1432 cttacggtgcaagatttggcagaagatcctcgcctcctcctcctcctcctcctcctcctcctcct 1491  
 QY 897 ccaagaagccaggttcccccgaatcctggtcttgagagaaaaaacaacactcggcgcaa 956  
 Db 1492 ccaagaagccaggttcccccgaatcctggtcttgagagaaaaaacaacactcggcgcaa 1551  
 QY 957 aaagtgcagccctcttctcgttatctaccccgagatcgagagccgagacttggccgga 1616  
 Db 1552 aaagtgcagccctc-tctcctatctaccccgagatcgagagccgagacttggccgga 1610  
 QY 1017 accccaactctgtcccccaccccttatccagacacaggtgctgtgagggga-cctctgc 1075  
 Db 1611 accccaactctgtcccccaccccttatccagacacaggtgctgtgagggagacccctcgc 1670  
 QY 1076 cctctctggagctcgtgtgtgagggacactgctgcgggacactcggagcggagagagcgc 1135  
 Db 1671 cctctctggagctcgtgtgtgagggacactgctgcgggacactcggagcggagggggcgc 1730  
 QY 1136 caccgggagcggagacagacagatcgcatattaccgctgcgcacactatggccctccat 1195  
 Db 1731 caccgggagcggagacagacagatcgcatattaccgctgcgcacactatggccctccac 1790  
 QY 1196 gcaaggggcaatttgagccctcctcagatttggcccttttctcagacatctctataa 1255  
 Db 1791 acggggggcgaattgagccctcctcagatttggcccttttctcagacatctctataa 1850  
 QY 1256 ttgaaacactaacacatcccttctcggagagatccccacgcctcacgggtgtgtgga 1315  
 Db 1851 ttgaaacactaacacatcccttctcggagagatccccacgcctcacgggtgtgtgga 1910  
 QY 1316 gtccttctgtctcaccagcctacttgggatgattgtcaacagctgctgcagacact 1375  
 Db 1911 gtccttctgtctcaccagcctacttgggatgattgtcaacagctgctgcagacact 1970  
 QY 1376 ctccaacacggagcggagagaaattctgttagaggtcagaaaaatgttctcctggggc 1435  
 Db 1971 ctccaacacggagcggagagaaattctgttagaggtcagaaaaatgttctcctggggc 2030  
 QY 1436 cgacggcgacccacacgagttgcaaatgagattgacatggatttcccttgcactcgc 1495  
 Db 2031 cgacggcgacccacacgagttgcaaatgagattgacatggatttcccttgcactcgc 2090  
 QY 1496 cgggttgggactcaacacggtgaaagtagggagagcttggaaatctatcgcagactct 1555  
 Db 2091 cgggttgggactcaacacggtgaaagtagggagagcttggaaatctatcgcagactct 2150  
 QY 1556 ggttggcggtctcggggcgccctcaagagcggccactaatttggctaaagtaagaggt 1615  
 Db 2151 ggttggcggtctcggggcgccctcaagagcggccactaatttggctaaagtaagaggt 2210  
 QY 1616 gatcaggggacccgaagacccctcctcgttattcttgagagctcatggaagcctcag 1675  
 Db 2211 gatcaggggacccgaagacccctcctcgttattcttgagagctcatggaagcctcag 2270  
 QY 1676 gcggttccaccccttttgcctcctcagagggcccgagagcctcagtgccctggcctt 1735



QY 3896 tcaaacgctgcttccacagagaggttcactttcgtccaccagcgctctcaaacctgc 3955  
Db 4491 tcaaacgctgcttccacagagaggttcactttcgtccaccagcgctctcaaacctgc 4550  
QY 3956 cactcttctgctgaagagactgatgaacacagtgactctatgattgcatcaactattgat 4015  
Db 4551 cactcttctgctgaagagactgatgaacacagtgactctatgattgcatcaactattgat 4610  
QY 4016 tgaggagactgggtccgaagacacttacacataaccgctgactggagagagtgctaac 4075  
Db 4611 tgaggagactgggtccgaagacacttacacataaccgctgactggagagagtgctaac 4670  
QY 4076 ctggttcactgacggaagcagctatgtgtggaagtaagagatgctgggcccgcgact 4135  
Db 4671 ctggttcactgacggaagcagctatgtgtggaagtaagagatgctgggcccgcgct 4730  
QY 4136 ggtgagcgggacccgcagatctgggcccagcagcctgcccgaagaaacttcgcgcacaaa 4195  
Db 4731 ggtgagcgggacccgcagatctgggcccagcagcctgcccgaagaaacttcgcgcacaaa 4790  
QY 4196 ggtgagcctatggccctcacgcaagctttgcgctggcgcgaaggaataccataaacat 4255  
Db 4791 ggtgagcctatggccctcacgcaagctttgcgctggcgcgaaggaataccataaacat 4850  
QY 4256 ttatacggacagcaggtatgctttgcgactgcacacgtacacggggccatctataaaca 4315  
Db 4851 ttatacagacagcaggtatgctttgcgactgcacacgtacacggggccatctataaaca 4910  
QY 4316 aagggggttgcttacctcgacggaggaggaataaagaacaaagaggaatcttaagcct 4375  
Db 4911 aagggggttgcttacctcgacggaggaggaataaagaacaaagaggaatcttaagcct 4970  
QY 4376 attagagccttacatttgcacaaagcctacgtattatacactgctcgtgacatcagaa 4435  
Db 4971 attagagccttacatttgcacaaagcctacgtattatacactgctcgtgacatcagaa 5030  
QY 4436 agccaaagatcctatctatagagggaacccagatgctgacgggttgcacagcagcagc 4495  
Db 5031 agccaaagatcctatctacagagggaacccagatgctgacgggttgcacagcagcagc 5090  
QY 4496 caagcgtgttaacctctcctatataagaacgcccacaaagcccagaaacccagacgaca 4555  
Db 5091 caagcgtgttaacctctcctatataagaacgcccacaaagcccagaaacccagacgaca 5150  
QY 4556 gtacacccctagagactgccaagagataaaaagatagaccagttctctgacatcccgga 4615  
Db 5151 gtacacccctagagactgccaagagataaaaagatagaccagttctctgacatcccgga 5210  
QY 4616 ggggacctgctatcacctatgggaaggaatacctgcccacaaagagggttagaata 4675  
Db 5211 agggacctgctatcacctatgggaaggaatacctgcccacaaagagggttagaata 5270  
QY 4676 tgtccaaagatacatcgtctaacccacccataggaactaaacacccctgcagcagttggtcag 4735  
Db 5271 tgtccaaagatacatcgtctaacccacccataggaactaaacacccctgcagcagttggtcag 5330  
QY 4736 aacatccctctatcatgttctgagcctaccagagtgctgactcgtggtggtcaaacattg 4795  
Db 5331 aacatccctctatcatgttctgagcctaccagagtgctgactcgtggtggtcaaacattg 5390  
QY 4796 tgtgacctgacagctggttaattgctaactcttccagaaacacccaggaagagactaacg 4855  
Db 5391 tgtgacctgacagctggttaattgctaactcttccagaaacacccaggaagagactaacg 5450  
QY 4856 ggggaagccacagcgctcactgggaagtgacttcaactgaggttaagccggttaataa 4915  
Db 5451 ggggaagccacagcgctcactgggaagtgacttcaactgaggttaagccggttaataa 5510  
QY 4916 cggaaacaaatctattgtttttgtagacacacccctttcagatggtgagagcttacc 4975  
Db 5511 cggaaacaaatctattgtttttgtagacacacccctttcagatggtgagagcttacc 5570  
QY 4976 tactaagaaagagacttcaacccgtgtgtgctaagaaataactgtggaggaaatttttccaaag 5035

Db 5571 tactaagaaagagacttcaacccgtgtgtgctaagaaataactgtggaggaaatttttccaaag 5630  
QY 5036 atttggataactaaggttaaggtgcagacaatgggtccagactttcgttgcacaggttaag 5095  
Db 5631 atttggataactaaggttaaggtgcagacaatgggtccagactttcgttgcacaggttaag 5690  
QY 5096 tcaggagactggccaagatattgggagattggaacctgattgcatcacagaccoca 5155  
Db 5691 tcaggagactggccaagatattgggagattggaacctgattgcatcacagaccoca 5750  
QY 5156 aagctcaggacaggttagagagatgaatagaaccattaaagagacccttaccacaaattgac 5215  
Db 5751 aagctcaggacaggttagagagatgaatagaaccattaaagagacccttaccacaaattgac 5810  
QY 5216 cacagagactggcatttaattgagtggtcctcctcgtcctcgttctgtgttttagggtag 5275  
Db 5811 cgcggagactggcgttaattgagtggtcctcctcgtcctcgttctgtgttttagggtag 5870  
QY 5276 gaacacccctggacagtttggcctgacccctatgaattgctctacggggagacccccc 5335  
Db 5871 gaacacccctggacagtttggcctgacccctatgaattgctctacggggagacccccc 5930  
QY 5336 gttggcagaataattgctttgcacatagtgctgattgctgctttccacagcctttgtctc 5395  
Db 5931 attggtagaataattgcttccgtacatagtgctgacgtgctgttccacagcctttgtctc 5990  
QY 5396 taggctcaagcgtcagtggtgggtgagcagcagcgtggaagcagctccggaggccta 5455  
Db 5991 taggctcaagcgtcagtggtgggtgagcagcagcgtggaagcagctccggaggccta 6050  
QY 5456 ctacaggagag 5466  
Db 6051 ctacaggagag 6061

RESULT 10  
AAC67032

ID AAC67032 standard; DNA; 6076 BP.

XX AAC67032;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 32.

XX Xenotransplantation; Infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian  
transplant recipients, by introducing recombinant DNA comprising DNA  
encoding extracellular proteins of the agent into donor cells, such as  
swine cells -

XX Claim 16; Page 117-119; 144pp; English.

XX The present invention provides a method to prevent the transmission of  
infectious agents during xenotransplantation. This involves introducing



QY 2036 caatagaccactcgaagaccagtggtggtattgtataagaaaaagacactgggc 2095  
DB 2631 caatagaccactcgaagaccagtggtggtattgtataagaaaaagacactgggc 2690  
QY 2096 aagaaactgcccgaaggaaggaacaaagaccgaaggttctagcttagaagataa 2155  
DB 2691 aagaaactgcccgaaggaaggaacaaagaccgaaggttctagcttagaagataa 2750  
QY 2156 agattagggagacgggttcggaccctcccggagcccgaggttaactttgaggtgga 2215  
DB 2751 agactaggaagacgggttcggaccctcccggagcccgaggttaactttgaggtgga 2810  
QY 2216 ggggaaccactgtagttctggtgataccggagggagcattcagttgctgataacc 2275  
DB 2811 ggggaaccactgtagttctggtgataccggagggagcattcagttgctgataacc 2870  
QY 2276 attaggaactaaaaaataaactctgggtgatgggtgcccaggggcaacggcagta 2335  
DB 2871 attaggaactaaaaaataaactctgggtgatgggtgcccaggggcaacggcagta 2930  
QY 2336 tccatgactaccgaagaccgttgaacttggagtgagtggaacgggttaaccactcgttct 2395  
DB 2931 tccatgactaccgaagaccgttgaacttggagtgagtggaacgggttaaccactcgttct 2990  
QY 2396 ggtcactcctgagtgcccagtaacccttctaggtagagacttactgaccaagatgggagc 2455  
DB 2991 ggtcactcctgagtgcccagtaacccttctaggtagagacttactgaccaagatgggagc 3050  
QY 2456 tcaaatctcttgaacaaaggaacagaaagtgtctgtgaaataacaaacccactactgt 2515  
DB 3051 acaaatctcttgaacaaaggaacagaaagtgtctgtgaaataacaaacccactactgt 3110  
QY 2516 gttagaccctcaattagatgaataatcagactatattcctcccaagtaagactataca 2575  
DB 3111 gttagaccctcaattagatgaataatcagactatattcctcccaagtaagactataca 3170  
QY 2576 agatatacagtcctgttgagcaggttcccgaagcctggcgagaacccgagggatggg 2635  
DB 3171 aaatatacaattctgttggaacagtttcccgaagcctggcgagaacccgagggatggg 3230  
QY 2636 ttgggaagcaaggttcccgaagcaggttattcaactgaagccagtgctacacagatac 2695  
DB 3231 ttgggaagcaaggttcccgaagcaggttattcaactgaagccagtgctacacagatac 3290  
QY 2696 agtcagacagtaacctttagtagagagctcgaagaaatgtgcccagatgttcaag 2755  
DB 3291 agtcagacagtaacctttagtagagagctcgaagaaatgtgcccagatgttcaag 3350  
QY 2756 attaatccaaagggcactcctgttcccaatccccttgaataactcccctgctacc 2815  
DB 3351 attaatccaaagggcactcctgttcccaatccccttgaataactcccctgctacc 3410  
QY 2816 ggttagaagcctgggaacaaatgattatcgaccagtagagacttagagaggtcaataa 2875  
DB 3411 ggttagaagcctgggaacaaatgattatcgaccagtagagacttagagaggtcaataa 3470  
QY 2876 aagggtgcaggacatacacccaagctccgaaccccttataacctcttgagcgcctccc 2935  
DB 3471 acgggtgcaggatacacccaagctccgaaccccttataacctcttgagcgcctccc 3530  
QY 2936 gctgaacgaactgttacacagatttgacttaaaagatgctcttcttctgctgagatt 2995  
DB 3531 accccaagagctggtatacagatttgacttaaaagatgctcttcttctgctgagatt 3590  
QY 2996 acacccactagccaaccacttttgccttcgaatggagatccagggtacgggaagac 3055  
DB 3591 acacccactagccaaccacttttgccttcgaatggagatccagggtacgggaagac 3650  
QY 3056 cgggagctcaactgacccgactgcccgaaggttcaagaactcccagacatcttga 3115  
DB 3651 cgggagctcaactgacccgactgcccgaaggttcaagaactcccagacatcttga 3710

QY 3116 cgaagcctacacaggacacttgccaaactcaggatccaacaccctcaggtgaacctct 3175  
DB 3711 cgaagcctacacaggacacttgccaaactcaggatccaacaccctcaggtgaacctct 3770  
QY 3176 ccagtacgtggatgacctgtcttggcgggagccaccacacaggaactgttagaaggtac 3235  
DB 3771 ccagtacgtggatgacctgtcttggcgggagccaccacacaggaactgttagaaggtac 3830  
QY 3236 gaaggcactactctgtggaattgtgacctaggtctacagagcctctgtaagaagccca 3295  
DB 3831 gaaggcactactctgtggaattgtgacctaggtctacagagcctctgtaagaagccca 3890  
QY 3296 gatttgacgagagaggttaacatacttgggttacagtttgcggggcgagcagatgct 3355  
DB 3891 gatttgacgagagaggttaacatacttgggttacagtttgcggggcgagcagatgct 3950  
QY 3356 gacggaggcacggaagaaaactgtagtccagataccggccccaacacacagcaaaagt 3415  
DB 3951 gacggaggcacggaagaaaactgtagtccagataccggccccaacacacagcaaaagt 4010  
QY 3416 gagaagtttttgggacagctggattttgcagactgtggtatcccggttttgacactt 3475  
DB 4011 gagaagtttttgggacagctggattttgcagactgtggtatcccggttttgacactt 4070  
QY 3476 agcagcccactctacccgttaacaaagaaaggggattctcctgggtcctgagca 3535  
DB 4071 agcagcccactctacccgttaacaaagaaaggggattctcctgggtcctgagca 4130  
QY 3536 ccagaagcatttgatgctatacaaaagccctgtgtagcgcaactgtcttgccctccc 3595  
DB 4131 ccagaagcatttgatgctatacaaaagccctgtgtagcgcaactgtcttgccctccc 4190  
QY 3596 tgaactaactaaacctttacccctttatgtgtagcgtaagggatgtagccgagagt 3655  
DB 4191 tgaactaactaaacctttacccctttatgtgtagcgtaagggatgtagccgagagt 4250  
QY 3656 tttaacccaaacccctagaccatgagagacactgttgcctacctgtcaagaagcttga 3715  
DB 4251 tttaacccaaacccctagaccatgagagacactgttgcctacctgtcaagaagcttga 4310  
QY 3716 tccgtagccagtggttgcccggtatgtctgaaggtatcgagctgtggccatactggt 3775  
DB 4311 tccgtagccagtggttgcccggtatgtctgaaggtatcgagctgtggccatactggt 4370  
QY 3776 caagcagctgacaacttgaacttgggacaaataaactgtatagccccactgactt 3835  
DB 4371 caagcagctgacaacttgaacttgggacaaataaactgtatagccccactgactt 4430  
QY 3836 ggagaacactgttgcgagcccccagacccgatggatgacacaccccgatgacccacta 3895  
DB 4431 ggagaacactgttgcgagcccccagacccgatggatgacacaccccgatgacccacta 4490  
QY 3896 tcaagcctgtcttcaagagaggggtcaacttgcctccacagcgcgtctcaacctgc 3955  
DB 4491 tcaagcctgtcttcaagagaggggtcaacttgcctccacagcgcgtctcaacctgc 4550  
QY 3956 cactcttgcctgaagactgtgaacccagtaactatgactatgactatgactatgact 4015  
DB 4551 cactcttgcctgaagactgtgaacccagtaactatgactatgactatgactatgact 4610  
QY 4016 tgaggagactgggtcccgaagacacttacagataccgctgactggagagtactaac 4075  
DB 4611 tgaggagactgggtcccgaagacacttacagataccgctgactggagagtactaac 4670  
QY 4076 ctggttactgacggaagcagctatgtgtgaaaggttaagaggtgctggggcgagct 4135  
DB 4671 ctggttactgacggaagcagctatgtgtgaaaggttaagaggtgctggggcgagct 4730  
QY 4136 ggtgagcgggcccgcagatctgtggccagcagctgcgcgaagaaacttcagcgcaaaa 4195  
DB 4731 ggtgagcgggcccgcagatctgtggccagcagctgcgcgaagaaacttcagcgcaaaa 4790  
QY 4196 ggtgagctatggccctcagcaagcttgcggctgcgcgaaggaataccataacat 4255



CC histological localisation of activated retrovirus. Using Polymerase Chain  
CC Reaction DNA Quantitation (PDQ) on blood mononuclear cells, infectivity  
CC titration and susceptibility testing can be performed. Ultimately animal  
CC donors without intact porcine retroviral sequences or with a lower copy  
CC number of viral elements could be selected.

SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

Query Match	60.7%;	Score 4986;	DB 18;	Length 8060;
Best Local Similarity	95.28;	Pred. No. 0;		
Matches 5276; Conservative	0;	Mismatches 245;	Indels 23;	Gaps 12;

Qy 1 gtggtgtacgactgtggggcccccagcgcgcttggaataaaaaatcctcttqctqtttqcac 60

[illegible][illegible]

61 aagaccgcttctcgtgagtgatt-tgggggtgcgcctcttccgagcccgacgagggga 119  
QY

Db 2590 aagaccgcttctcgtgagtgaataaggggagtcgccttttccagacctggagq----- 2643

[illegible]

17 18 19

Db 2643 --ttcttttttgctggtctttacatttggggctcgccgggatctgtcgcgccacccta 2700

Qy 180 acaccgagaccgacttgaggta-aaqqgataccctttqqaacqtatata 236

db 2701

2701 acacccyyaaccyyagggcdaaaggatccctctctcttaacgtgtatgcatgtacc 2760

QY 237 ggccggcgtctctgttcttgagtgctctgttttcgggtgatgcgcgcttctcggttgcagctg 296

Db 2761 ggccggcgctctgtcttgagtggtctgttttccagtgaatcgcgcctttccagatttcacagctg 2820

[illegible]

QY zzz' tccccagaccglaaggactggaggactgtgatcagcagacgtgctaggaggatcacag 356  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2821 tccttcaggccgtaaggctgggggactgtgatcagcagacgtgctaggaggtcacag 2880

Qy 357 gctgccacctgggggacqccccqggaggtgaggagagccaggagagccctggtggtctcc 416

2001

50 z88r yclcyccgcgggacgcccccggagggtgaggtgagcagcgacctgggtctcc 2940

Qy 417 tactgtcgggtcagaggaccgagttctgttgttgaagcgaaagcttccccctccgcccg 476

Db 2941 tactgtcgggtcagaggaaccqaattctgtttgctgaagcgaagccttccccctccgcgacg 3000

[illegible]

QY 477 tccgactctcttgcctgcttggaagacgcggacgggtcgcgtgtgtctggatctgttgcg 536

Db 3001 tccgactcttttgccctgttggaatacgtggacgggtcacgtgtgtctgatatctgtg 3060

Qy 537 gtttctgtctcgtgtgtcttttgtcttgtatgcgtcccttgtctacagttttaatataggacag 506

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

DB 3061 gtttctgttttctgtgtctcttctgtgtgtccttgtctctacagttttaatatgggacag 3120

QY 597 acagtgactaccccccttagtttgactctcgaccattggactgaagttagatccagggt 656

Db 3121 acggtgacgacccccctcttaatttgaactctcgaccattgaactaaacttaaatccacggct 3180

110



Db 4500 ggaagagagagagatttttaggaaataggctcaggccctagacagtcagggaacctggg 4559  
Qy 2036 caataggaccactcgcacaggaccagtgctgattgttaaaagaaagacactgggc 2095  
Db 4560 caataggaccactcgcacaggaccagtgctgattgttaaaagaaagacactgggc 4619  
Qy 2096 aaggactgcccagaagaggaacaaagagaccgaagtgctcctagagaagataa 2155  
Db 4620 aaggactgcccagaagaggaacaaagagaccgaagtgctcctagagaagataa 4679  
Qy 2156 agattaggaggagacgggttgcgaccctcccgagcccgaggggttaactttgaagggtga 2215  
Db 4680 agattaggaggagacgggttgcgaccctcccgagcccgaggggttaactttgaagggtga 4739  
Qy 2216 gggcaacacagtgacttccctggttgatcccgagcagagcattcagtgctcagacac 2275  
Db 4740 gggcaacacagtgacttccctggttgatcccgagcagagcattcagtgctcagacac 4799  
Qy 2276 attagaaaactaaagaaaataatcctgggtgatgggtgcccagggcagcagta 2335  
Db 4800 attagaaaactaaagaaaataatcctgggtgatgggtgcccagggcagcagta 4858  
Qy 2336 tccatggactaccgaagacacttgacttggagtgaggcgggttaacccactcgtttct 2395  
Db 4859 tccatggactaccgaag-acagttgacttggagtgaggcgggttaacccactcgtttct 4917  
Qy 2396 ggtcaccctgagtgcccagtcacccttctaggtagagacttactgacacagatggagc 2455  
Db 4918 ggtcaccctgagtgcccagtcacccttctaggtagagacttactgacacagatggagc 4977  
Qy 2456 tcaaatcttttgacaggaagaccagaagtgctgtgataacaaacccatcactgt 2515  
Db 4978 tcaaatcttttgacaggaagaccagaagtgctgtgataacaaacccatcactgt 5037  
Qy 2516 gttgacctccaattagatgatgaatatcgactatatctcccaagtaaaagcctgatca 2575  
Db 5038 gttgacctccaattagatgatgaatatcgactatatctccctccctagtaaaagcctgatca 5097  
Qy 2576 agatatacagtgctgttgagcaggttccccaagcctggcagaaaacccaggaagg 2635  
Db 5098 aaatatacaattcgttggaacagttccccaagcctggcagaaaacccaggaagg 5157  
Qy 2636 tttggcaagcaagttccccaagcctggcagaaaacccaggaagg 2695  
Db 5158 tttggcaagcaagttccccaagcctggcagaaaacccaggaagg 5217  
Qy 2696 agtcagacagtcacccttgagtagagagcctcgagaagaaattggccgcatgtcaag 2755  
Db 5218 agtcagacagtcacccttgagtagagagcctcgagaagaaattggccgcatgtcaag 5277  
Qy 2756 attaatcccaaggcactcctagttcctgtccaatcccttggaaatactccctgctacc 2815  
Db 5278 attaatcccaaggcactcctagttcctgtccaatcccttggaaatactccctgctacc 5337  
Qy 2816 ggttaggaagcctgggaccacaaatgattatagaccagtagagagcttgagagggtaataa 2875  
Db 5338 ggttaggaagcctgggagcactaagtactatgacacagtagagagcttgagagggtaataa 5397  
Qy 2876 aagggtgagagacatacaccccaagctccggaacccctataacotcttgagcgcctccc 2935  
Db 5398 aagggtgagagacatacaccccaagctccggaacccctataacotcttgagcgcctccc 5457  
Qy 2936 gctgaacaggaaactggtcacacagttattggaacttaaaagatgcttctctcctgagatt 2995  
Db 5458 accccaacaggagctggttacacagttattggaacttaaaagatgcttctctcctgagatt 5517  
Qy 2996 aacccccactagcccaacacacttttgccttcgaattggagagatccaggtacgggaagac 3055  
Db 5518 aacccccactagcccaacacacttttgccttcgaattggagagatccaggtacgggaagac 5577  
Qy 3056 cggggcagctcacctggaccacactgcccacagggttcaagaaactcccgaccacttttga 3115

Db 5578 cgggcagctcacctggaccgactgcccgaagggttcaagaaactcccgaccacttttga 5637  
Qy 3116 cgaagccctacacagagactgcccgaacttcaggatcccaacccctcaggtgacctct 3175  
Db 5638 cgaagccctacacagagactgcccgaacttcaggatcccaacccctcaggtgacctct 5697  
Qy 3176 ccagtcagtgagtgactgcttctgctgggagcccaaacacaggaactgcttagaaggta 3235  
Db 5698 ccagtcagtgagtgactgcttctgctgggagcccaaacacaggaactgcttagaaggta 5757  
Qy 3236 gaaggcaactactgctggaattgtctgacctaggctacagagcctctgctaaagagccca 3295  
Db 5758 gaaggcaactactgctggaattgtctgacctaggctacagagcctctgctaaagagccca 5817  
Qy 3296 gatttcagagagagagtaacatacttgggtgaggttgcggggcggcagcgatggct 3355  
Db 5818 gatttcagagagagagtaacatacttgggtgaggttgcggggcggcagcgatggct 5877  
Qy 3356 gacggaggcaggaagaaactagtccaagataccgggcccccaaacacagcccaacaagt 3415  
Db 5878 gacggaggcaggaagaaactagtccaagataccgggcccccaaacacagcccaacaagt 5937  
Qy 3416 gacagagtttttgagacagctgatttgcagactgtgatcccggttttgcgacct 3475  
Db 5938 gacagagtttttgagacagctgatttgcagactgtgatcccggttttgcgacct 5997  
Qy 3476 agcagccccacttaccgctaaccaagaaaggggattctctgggtcctcagca 3535  
Db 5998 agcagccccacttaccgctaaccaagaaaggggattctctgggtcctcagca 6057  
Qy 3536 ccaggaagcatttgatgctatacaaaagccctgctgagcgaacctgctctggccctccc 3595  
Db 6058 ccaggaagcatttgatgctatacaaaagccctgctgagcgaacctgctctggccctccc 6117  
Qy 3596 tgacgtaactaaacctttaccctttatgtggatgagctgaaggagtagcccgagaggt 3655  
Db 6118 tgacgtaactaaacctttaccctttatgtggatgagctgaaggagtagcccgagaggt 6177  
Qy 3656 tttaacccaaacctagagaccatggagagacactgttgcctactgtcaaaagagttga 3715  
Db 6178 tttaacccaaacctagagaccatggagagacactgttgcctactgtcaaaagagttga 6237  
Qy 3716 tccgttagcagtggttgccctgtatgtgaaggctatcgcagctgctgagcactatggt 3775  
Db 6238 tccgttagcagtggttgccctgtatgtgaaggctatcgcagctgctgagcactatggt 6297  
Qy 3776 caagagcgtgacaaattgactttgggac-agaataaactgtaataagcccccatgcat 3834  
Db 6298 caagagcgtgacaaattgactttgggacagaagaataaactgtaataagcccccatgcat 6357  
Qy 3835 tgagaaacatcgttgcggcagccccagacagatggatgacacaaaccccgatgacct 3894  
Db 6358 tgagaaacatcgttgcggcagccccagacagatggatgacacaaaccccgatgacct 6417  
Qy 3895 atcaaacctgcttctcacagagaggtcacttgcctccacacagcgcctcaacccctg 3954  
Db 6418 atcaaacctgcttctcacagagaggtcacttgcctccacacacagcgcctcaacccctg 6477  
Qy 3955 ccactcttctgctgaagagactgataccagtgactgactgactgactgactgactgactgact 4014  
Db 6478 ccactcttctgctgaagagactgataccagtgactgactgactgactgactgactgactgact 6537  
Qy 4015 ttgaggagactgggttcgcaagagacacttcacagacataccgctgactgaggaagtgttaa 4074  
Db 6538 ttgaggagactgggttcgcaagagacacttcacagacataccgctgactgaggaagtgttaa 6597  
Qy 4075 cctgttctactgacggaagcagctatgtgtgaaagtagagagtggtcggggcggcag 4134  
Db 6598 cctgttctactgacggaagcagctatgtgtgaaagtagagagtggtcggggcggcag 6657  
Qy 4135 tgggtgagcgggacccgacgatctggccagcagcctgcccgaaggaacttcagcgcaaa 4194  
Db 6658 tgggtgagcgggacccgacgatctggccagcagcctgcccgaaggaacttcagcgcaaa 6717





QY 1 ggggtgtacgactgtggcccgccgagcgcttggaataaaaaatcctcttctgtgtttgcac 60  
Db 2530 ggggtgtacgactgtggcccgccgagcgcttggaataaaaaatcctctgtgtttgcac 2589  
QY 61 aagaccgtctctgtgagtgatt-tgggggtgcgctcttcgagcccgagcgaggggga 119  
Db 2590 aagaccgtctctgtgagtgatttgggggagtcgctcttccgagcgtggag----- 2642  
QY 120 ttgtcttttactgoccttcaattgtgtgttgccggggaataactctgcgaccacccott 179  
Db 2643 --ttcttttctgtgtcttcaatttgggggctcgtcccggatctctgcgcccaccccta 2700  
QY 180 acaccgagaacgacttgagagta-aagggatccctcttggaaactgtgtg--tgtgtc 236  
Db 2701 acaccgagaacgacttgagagtaaaagatcctcttttaacgtgtatgcatgtacc 2760  
QY 237 ggcggcgctctctgttctgtgagtgctgttttgggtgtagcgcgcttctcggtttcagctg 296  
Db 2761 ggcggcgctctctgttctgtgagtgctgttttcaagtgtgtgcgcttctcggtttcagctg 2820  
QY 297 tctctcagacgttaagactgagagctgtgatcagcagacgtgtctagagagatcacag 356  
Db 2821 tctctcagacgttaagggctgggggactgtgatcagcagacgtgtctagagagatcacag 2880  
QY 357 gctgcacccctggggagcccgccggaggtggggagagccaggagcgcctgtgtgtctcc 416  
Db 2881 gctgtgcctggggagcccgccggaggtgagagagccaggagcgcctgtgtgtctcc 2940  
QY 417 tactgtcgttcagagacccgaattctgtgttgtaagcagaaagacttcccccctccggccg 476  
Db 2941 tactgtcgttcagagacccgaattctgtgttgtaagcagaaagacttcccccctccggccg 3000  
QY 477 tccgactctttgcctgtctgtggaagacggcgagcggtgcgctgtctgtgatctgttg 536  
Db 3001 tccgactctttgcctgtctgtggaatacgttgagcgggtcagcgtgtctgtgatctgttg 3060  
QY 537 gttctgtctgtgtgtctgttctgtgtgtgtctgtgtctgtgtctgtgtgtgtgtgtgt 596  
Db 3061 gttctgtgtgtgtgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3120  
QY 597 acagtgactacccctctagtgtactctcagccattgagctgaagttagatccaggct 656  
Db 3121 acgggtgacacccctctagtgtactctcagccattgagctgaagttagatccaggct 3180  
QY 657 cataattgtcaggttcaggttaagaaggacacctggcagaccttctgtgcctctgaaatgg 716  
Db 3181 cataattgtcaggttcaggttaagaaggacacctggcagaccttctgtgcctctgaaatgg 3240  
QY 717 ccaacattcagatgttgatggccatcagaggggacacctttaaattctgaaattatcctggct 776  
Db 3241 ccgacattcagatgttgatggccatcagaggggacacctttaaattctgaaattatcctggct 3300  
QY 777 gttaaagcaatcatttttcagactgagaccgctctcattctgtatcagagccctatatc 836  
Db 3301 gttaaagcaatcatttttcagactgagaccgctctcattctgtatcagagccctatatc 3360  
QY 837 ctacgtggaagatttggcagaagaatcccgccatgggttaaaccatggctgaaataaa 896  
Db 3361 ctacgtggaagatttggcagagatcctccgcatgggttaaaccatggctgaaataag 3420  
QY 897 ccaagaaacgaggtcccgaaatcctggctctgttgagagaaaaaacaacactcggccgaa 956  
Db 3421 ccaagaaacgaggtcccgaaatcctggctctgttgagagaaaaaacaacactcggctgaa 3480  
QY 957 aagtcagccctctctctgtatctacccgagatcgagagcgccgacacttggccgga 1016  
Db 3481 aagtcagccctc--tctcctatctaccccgagatcgagagcccgcttggccgga 3539  
QY 1017 accccaactgtttccccccaccccttatccagacaggggtgtgtgagggga--cctctgc 1075  
Db 3540 accccaactgtttccccccaccccttatctgcacaggggtgcgagggggaccccttgc 3599

QY 1076 cctctctggagctccggtgtgagagaaactcgtcgcgagactcgcgagccgagagagcgc 1135  
Db 3600 cctctctggagctccggtgtgagagaaactcgtcgcgagactcgcgagccgagagagcgc 3659  
QY 1136 caccggagcgcagacagagatcgcatattaccgctgcgcacatattggccctcccat 1195  
Db 3660 caccggagcgcagacagagatcgcatattaccgctgcgcacatattggccctcccat 3719  
QY 1196 gccagggggccaattgcagccctccagctattggccctttttcttcgcagatctctataa 1255  
Db 3720 accggggggccaattgcagccctccagctattggccctttttcttcgcagatctctataa 3779  
QY 1256 ttggaaaactaacatccccctctctcgcagagatcccccaacgctcacgggttggtaga 1315  
Db 3780 ttggaaaactaacatccccctctctcgcagagatcccccaacgctcacgggttggtaga 3839  
QY 1316 gtcccttatgtctctcaccagcctacttgggatgtgtcacaagctgctgcagacact 1375  
Db 3840 gtcccttatgtctctcaccagcctacttgggatgtgtcacaagctgctgcagacact 3899  
QY 1376 cttcacaccgagagcgagagaaattctgttagaggttagaaaaatgttctctgggccc 1435  
Db 3900 cttcacaccgagagcgagagaaattcttattagaggttagaaaaatgttctctgggccc 3959  
QY 1436 cgacggcgaccccaacgagittgcaaaatgagattgacatgggatttcccttgactcgc 1495  
Db 3960 cgacggcgaccccaacgagttgcaaaatgagattgacatgggatttcccttgactcgc 4019  
QY 1496 cgggtgggactacaacacgctgaagtagggagagcttgaaaatctatccgagggctct 1555  
Db 4020 cgggtgggactacaacacgctgaagtagggagagcttgaaaatctatccgagggctct 4079  
QY 1556 ggtggcggtgtctcggggcgctcgaagcggccacttaatttggctaaagtagagaggt 1615  
Db 4080 ggtggcggtgtctcggggcgctcgaagcggccacttaatttggctaaagtagagaggt 4139  
QY 1616 gatcagggacccaacgaacactcctcgtatttcttgagagctcatggaagccttcag 1675  
Db 4140 gatcagggacccaacgaacactcctcgtatttcttgagagctcttggaagccttcag 4199  
QY 1676 gcggtcacccctttgatcctcagagcccgccagaaagcctcagtgccctggcctt 1735  
Db 4200 gcggtcacccctttgatcctcagagcccgccagaaagcctcagtggtcttggcctt 4259  
QY 1736 cattggcgagtcggtctgtgatatcagaagaaacttcagagactggaaggtttacagga 1795  
Db 4260 tataggcagtcagccttgatattagaaagagcttcagagactggaaggtttacagga 4319  
QY 1796 ggcgtgagtacgtgatctagttagagagggcagagaaggtgtattacagaaaggagacaga 1855  
Db 4320 ggcgtgagtacgtgatctagttagagagggcagagaaggtgtattacaaaggagacaga 4379  
QY 1856 agaggaagaagacagagaaaaagggagagagaagaagggaggaagcgtgtag 1915  
Db 4380 agaagaagggaacaaagaaaaagagagaagaagggaggaagacgtaataa 4439  
QY 1916 acggcagaagaagaatttgaactaagatcttggccgaggtgttgaaaggagagacagcag 1975  
Db 4440 acggcagaagaagaatttgaactaagatcttggctgcgaggtgttgaaagggaagcaaac 4499  
QY 1976 ggagagagagagatatttaggaaaaatttagtcaggccctcagacagtcagggaacccctggg 2035  
Db 4500 ggaaagagagagagatatttaggaaaaatttagtcaggccctcagacagtcagggaacccctggg 4559  
QY 2036 caatagaccctcctcagacaggaccagctgtgcgtattgtaagaaaaaaggaacactgggc 2095  
Db 4560 caatagaccctcctcagacaggaccagctgtgcgtattgtaagaaaaaaggaacactgggc 4619  
QY 2096 aaggaactgcacaaaggggaacaaagacagcagaggttctagctctagaagaagataa 2155  
Db 4620 aaggaactgcacaaaggggaacaaagacagcagaggttctagctctagaagaagataa 4679  
QY 2156 agattaggggagagcgggttctcggaacccctcctcccgagcccgaggttaacttgaaggtgga 2215

|||||  
Db 4680 agattagggagacgggttcgagcccccctcccgagccagggttaactttgaagtgga 4739  
QY 2216 ggggcaaccagttgagttcctggttgataccggagcggagcattcaagtctgtacaacc 2275  
Db 4740 ggggcaaccagttgagttcctggttgataccggagcggagcattcaagtctgtacaagc 4799  
QY 2276 attaggaataaaagaaaatacctgggtgatgggtgcccacagggcaacggcagta 2335  
Db 4800 attaggaataaaagaaaatacctgggtgatgggtgcccacagggcaacggcagta 4858  
QY 2336 tccatggactaccgaagaacccgttgactggagtgggagcgggtaaacccactcgtttct 2395  
Db 4859 tccatggactaccgaag - acagttgacttggagtgggagcgggtaaacccactcgtttct 4917  
QY 2396 ggtcatccctgagtgccagtagaccctcttagtagagacttactgaccaagatggagc 2455  
Db 4918 ggtcatccctgagtgccagtagaccctcttagtagagacttactgaccaagatggagc 4977  
QY 2456 tcaaatctttgaaacaaggaagaccagaagtgtctgtaataacaaacccatcactgt 2515  
Db 4978 acaaatctttgaaacaaggaagaccagaagtgtctgtaataacaaacccatcactgt 5037  
QY 2516 gttgacctccaaatagatgaatatogactatatctccccaagtaaaagcctgatac 2575  
Db 5038 gttgacctccaaatagatgaatatogactatatctccccaagtaaaagcctgatac 5097  
QY 2576 agatatacagttcctggttgagcagtttcccaagcctggcgagaacccgagggatggg 2635  
Db 5098 aatatacaatcttgggttgaaacagtttcccaagcctggcgagaacccgagggatggg 5157  
QY 2636 ttgtgcaagcaagttcccccacaggttatccaactgaagccagtgctacacaggtatc 2695  
Db 5158 ttgtgcaagcaagttcccccacaggttatccaactgaagccagtgctacacaggtatc 5217  
QY 2696 agtcagacgtaccccttgtagtagaggtctcgagaggaatttggcccgatgttcaag 2755  
Db 5218 agtcagacgtaccccttgtagtagaggtctcgagaggaatttggcccgatgttcaag 5277  
QY 2756 attaatccaagggatccctagttcctgtccaatcccttggaatacctccctgtacc 2815  
Db 5278 attaatccaagggatccctagttcctgtccaatcccttggaatacctccctgtacc 5337  
QY 2816 ggttagaagcctgggaccaatgatattcgaccagtagcacgaggttgagaggtcaataa 2875  
Db 5338 ggttagaagcctgggaccaatgatattcgaccagtagcacgaggttgagaggtcaataa 5397  
QY 2876 aaggttgaggaacatacacccacggtcccgaaacccctataacccctttgagggccctcc 2935  
Db 5398 acgggtgacaggaatacacccacaggtcccgaaacccctataacccctttggtgctctcc 5457  
QY 2936 gctggaacggaactggttacacagatttgacttaaaagatgctcttctgtcgtgagatt 2995  
Db 5458 accccaagagctggtatacagatttgacttaaaagatgctcttctgtcgtgagatt 5517  
QY 2996 acaccccaactagccaaccccttttgcctcgaaatggagatccagggtacgggaagaac 3055  
Db 5518 acaccccaactagccaaccccttttgcctcgaaatggagatccagggtacgggaagaac 5577  
QY 3056 cgggagctcaactgagccagctgcccgaaggttcaagaactccccagacatctttga 3115  
Db 5578 cgggagctcaactgagccagctgcccgaaggttcaagaactccccagacatctttga 5637  
QY 3116 cgaagccctacacaggaactggccaacttcaggatcccaacccctcaggtgacctctct 3175  
Db 5638 cgaagccctacacaggaactggccaacttcaggatcccaacccctcaggtgacctctct 5697  
QY 3176 ccagtagctggatgactgcttctgcgggagcccaacacaggaactgcttagaaggtac 3235  
Db 5698 ccagtagctggatgactgcttctgcgggagcccaacacaggaactgcttagaaggtac 5757  
QY 3236 gaaggcaactactggaatttctgacgtaggtctacagagcctctgtataaaggccca 3295  
|||||

Db 5758 gaaggcaactactgctggaatttctgtaacctaggtctacagagcctctgtataaaggccca 5817  
QY 3296 gatttcaggaagagaggttaacatacttgggttacagtttgcggggcgagcagatgggt 3355  
Db 5818 gatttcaggaagagaggttaacatacttgggttacagtttgcggggcgagcagatgggt 5877  
QY 3356 gacggaggcacgaagaaaacitgtagtcagataccgcccacacagcccaacaaagt 3415  
Db 5878 gacggaggcacgaagaaaacitgtagtcagataccgcccacacagcccaacaaagt 5937  
QY 3416 gagagagtttttggggacagctggattttgcagactgtgatacccggttttgcacctt 3475  
Db 5938 gagagagtttttggggacagctggattttgcagactgtgatacccggttttgcacctt 5997  
QY 3476 agcagcccaactctacccgttaacaaagaaaaggggattctcctgggtccctgagca 3535  
Db 5998 agcagcccaactctacccgttaacaaagaaaaggggattctcctgggtccctgagca 6057  
QY 3536 ccagaagggcatittgatgtatcaaaaagccctgtgagcgaactgtctgtggccctccc 3595  
Db 6058 ccagaagggcatittgatgtatcaaaaagccctgtgagcgaactgtctgtggccctccc 6117  
QY 3596 tgacgttaactaaacccctttacccctttatgtgagtgagcgttaaggagtagcccgaggagt 3655  
Db 6118 tgacgttaactaaacccctttacccctttatgtgagtgagcgttaaggagtagcccgaggagt 6177  
QY 3656 tttaaccccaacccctaggaactgagagagacccctgttgcctacactgttcaagaagacttga 3715  
Db 6178 tttaaccccaacccctaggaactgagagagacccctgttgcctacactgttcaagaagacttga 6237  
QY 3716 tctgtagccagtggttggcccgatgtctgaagcgtatcgagcgttggccatactggt 3775  
Db 6238 tctgtagccagtggttggcccgatgtctgaagcgtatcgagcgttggccatactggt 6297  
QY 3776 caaggacgtgacaattgaactttgggac - agaataactgtataagccccccatgact 3834  
Db 6298 caaggacgtgacaattgaactttgggacaagaataaactgtataagccccccatgact 6357  
QY 3835 tgaggaacatcgttcggcagccccccagaccgagatgagataccacgcccgcagatgacct 3894  
Db 6358 tgaggaacatcgttcggcagccccccagaccgagatgagataccacgcccgcagatgacct 6417  
QY 3895 atcaagcctgcttctcacagaggggtcaacttctgcctccaccagcgtctccaacctg 3954  
Db 6418 atcaagcctgcttctcacagaggggtcaagttcgttccaccacagcgtctccaacctg 6477  
QY 3955 ccactcttctgctgaagagactgatgaaccagtgactcaatgattgccatcaactatga 4014  
Db 6478 ccactcttctgctgaagagactgatgaaccagtgactcaatgattgccatcaactatga 6537  
QY 4015 ttgaggagactgggtccgcaagacccctacagacatcccgtagctgagagatgctaa 4074  
Db 6538 ttgaggagactgggtccgcaagacccctacagacatcccgtagctgagagatgctaa 6597  
QY 4075 cctggttcaactgacggaagcagctatgtggtggaaggttaagagatggttggcgaggag 4134  
Db 6598 cctggttcaactgacggaagcagctatgtggtggaaggttaagagatggttggcgaggag 6657  
QY 4135 tggtagagcggaccgcagcatctgggcagcagcctcccggaaggaactcaogcgaaa 4194  
Db 6658 tggtagagcggaccgcagcatctgggcagcagcctcccggaaggaactcaogcgaaa 6717  
QY 4195 aggcgtgagctatggccctcagcaagcttgcggctggcggaagggaaaatccataaca 4254  
Db 6718 aggcgtgagctatggccctcagcaagcttgcggctggcggaagggaaaatccataaca 6777  
QY 4255 tttaacgggacagcaggtatgcctttgcagctgcacacgttacacggggccctctataaac 4314  
Db 6778 tttaacgggacagcaggtatgcctttgcagctgcacacgttacacggggccctctataaac 6837  
QY 4315 aaagggggttgccttacctcagcagggagggaaaataaagaacaaagagaaattctaagcc 4374  
Db 6838 aaagggggttgccttacctcagcagggagggaaaataaagaacaaagagaaattctaagcc 6897



QY 2180 cccctcccgagcccggttaactttgaagtgagggcaaccagttcagttccctggt 2239  
Db 241 cccctcccgagcccggttaactttgaagtgagggcaaccagttcagttccctggt 300  
QY 2240 tgataccggagcgagcattcagtgctgctacaaccattaggaacacataaaagaaaaa 2299  
Db 301 tgataccggagcaaacattcagtgctactacagccattaggaacacataaaagaaaa 360  
QY 2300 atcctgggtgagtggtgcccacagggcaacggcagatccatggactaccggcaagaaacgt 2359  
Db 361 atcctgggtgagtggtgcccacagggcaacacagatccatggactaccggcaagaaacgt 420  
QY 2360 tgactgggagtgaggacgggtaacccactggtttctggtcatccctgagtgccagtaacc 2419  
Db 421 tgactgggagtgaggacgggtaacccactggtttctggtcatccctgagtgccagtaacc 480  
QY 2420 cctctaggtgagacttactgaccagatgggagctcaaatcttcttgaacaaagaaag 2479  
Db 481 cctctaggtgagacttactgaccagatgggagctcaaatcttcttgaacaaagaaag 540  
QY 2480 accagagtgctgtgataaacaacacatcaactggttgacccctcaattagatgata 2539  
Db 541 accagagtgctgtgataaacaacacatcaactggttgacccctcaattagatgata 600  
QY 2540 atatcgactatattctcccaagtaaacgctgatacaagatatagctcctggttgagca 2599  
Db 601 atatcgactatattctccctagtaaacgctgatacaagatatagctcctggttgagca 660  
QY 2600 gtttcccaagcctgggcagaaacgcaggatgggttggcgaagcaagtcccccaca 2659  
Db 661 gtttcccaagcctgggcagaaacgcaggatgggttggcgaagcaagtcccccaca 720  
QY 2660 gtttattcaactgaagccagtgataccagatcaactgagtcagacagtaacccttgagtag 2719  
Db 721 agttattcaactgaagccagtgccacacagtgccagtgatgagcagtaacccttgagtaa 780  
QY 2720 agagctcgagaagaaattggccgcatgttcaagatttaacacagggcatcctcagt 2779  
Db 781 agaagctcaagaagaaattggccgcatgtccaaagatttaacacagggcatcctcagt 840  
QY 2780 tccgtccaatcccttggaatactccctgctaccggttaggaagcctgggaccaatga 2839  
Db 841 tccgtccaatcccttggaatactccctgctaccggttaggaagcctgggaccaatga 900  
QY 2840 ttatcgaccagtcagagactgagaggtcaataaaaggtgagacatacacaccaac 2899  
Db 901 ttatcgaccagtcagagactgagaggtgcaataaaaggtgagagatacacaccaac 960  
QY 2900 ggtccgaaccttataacctctgagcgccctcccgctgaaacggaaactggtacacagt 2959  
Db 961 agtccgaaccttataacctctgagcgccctcccgctgaaacggaaactggtacacagt 1020  
QY 2960 attggaactaaagatgctctctgctgagattacacccactagccaacactttt 3019  
Db 1021 attggaactaaagatgctctctgctgagattacacccactagccaacactttt 1080  
QY 3020 tgccttcgaatggagaagatccagtagcgaagaacggcgagctcactgagaccact 3079  
Db 1081 tgccttcgaatggagaagatccagtagcgaagaacggcgagctcactgagaccact 1140  
QY 3080 gcccaaggttcaagaactcccgacctcttgaagaacccctacacagggacctggc 3139  
Db 1141 gcccaaggttcaagaactcccgacctcttgaagaacccctacacagagacctggc 1200  
QY 3140 caacttcaggtatccaaacctcaggtgacccctcctccagtaactgagacctgtctct 3199  
Db 1201 caacttcaggtatccaaacctcaggtgacccctcctccagtaactgagacctgtctct 1260  
QY 3200 ggcgggagcccaacaaagagctgcttagaaggtacgaagcactactgctggaatgtc 3259  
Db 1261 ggcgggagcccaacaaagagctgcttagaaggtacgaagcactactgctggaatgtc 1320  
QY 3260 tgacctaggtcacagagcctctgctgaagaagcccagatttgagagagaggttaacata 3319

Db 1321 tgacctaggtcacagagcctctgctgaagaagcccagatttgagagagaggttaacata 1380  
QY 3320 cttggggttacagtttgcggggcgagcgatggctgacggagggcagcaaaacactgt 3379  
Db 1381 cttggggttacagtttgcggggcgagcgatggctgacggagggcagcaaaacactgt 1440  
QY 3380 agtccagatcacggcccaacacacacacaaagtgagagagtttttgggacagctgg 3439  
Db 1441 agtccagatcacggcccaacacacacacaaagtgagagagtttttgggacagctgg 1500  
QY 3440 attttgcagacttgatcccggtttgcaccttagcagcccaactataccctaac 3499  
Db 1501 attttgcagacttgatcccggtttgcaccttagcagcccaactataccctaac 1560  
QY 3500 caaagaaaagggggtattctcctggtcctgagcaccagaaggaatttgatgctatcaa 3559  
Db 1561 caaagaaaagggggtattctcctggtcctgagcaccagaaggaatttgatgctatcaa 1620  
QY 3560 aaagggcctgctgagcgacactgctgcccctccctgacgtaaactaaacccctttaccct 3619  
Db 1621 aaagggcctgctgagcgacactgctgcccctccctgacgtaaactaaacccctttaccct 1680  
QY 3620 ttatgtggtgagcgtaaaggagtagcccgaggagttttaaccacaaacccctaggaccatg 3679  
Db 1681 ttatgtggtgagcgtaaaggagtagcccgaggagttttaaccacaaacccctaggaccatg 1740  
QY 3680 gaaggaacctgttgctactcctgcaaaagagcttgatcctgtagcagtggttgccct 3739  
Db 1741 gaaggaacctgttgctactcctgcaaaagagcttgatcctgtagcagtggttgccct 1800  
QY 3740 atgtctgaaggtctatcgacgtggtccatactggtcaagggacgtgacaaattgacttt 3799  
Db 1801 atgtctgaaggtctatcgacgtggtccatactggtcaagggacgtgacaaattgacttt 1860  
QY 3800 gggacagaaataactgttaattagcccccatgcatggtggaacatcgttgcgagcccc 3859  
Db 1861 gggacagaaataactgttaattagcccccatgcatggtggaacatcgttgcgagcccc 1920  
QY 3860 agaccagtgatgacaaacgcgcgcacccatgacccatcaaaagccttctcacagagag 3919  
Db 1921 agaccagtgatgacaaacgcgcgcacccatgacccatcaaaagccttctcacagagag 1980  
QY 3920 ggtcaacttgcctccacagcgcctcctcaacctgcaactcttctcctggaagagactga 3979  
Db 1981 ggtcaacttgcctccacagcgcctcctcaacctgcaactcttctcctggaagagactga 2040  
QY 3980 tgaacagtgactcatgattgcccataactattgattgagagagactgggtccgcaagga 4039  
Db 2041 tgaacagtgactcatgattgcccataactattgattgagagagactgggtccgcaagga 2100  
QY 4040 ccttacagacataccgctgactgagaaagttaacctggttcaactgacggaagcagcta 4099  
Db 2101 ccttacagacataccgctgactgagaaagttaacctggttcaactgacggaagcagcta 2160  
QY 4100 tgtgtgtgaaagtgaaaggtggtggcgagtggtggaacggacccgcacgactgtg 4159  
Db 2161 tgtgtgtgaaagtgaaaggtggtggcgagtggtggaacggacccgcacgactgtg 2220  
QY 4160 ggcagcagcctccggaaagaaacttcagcgcaaaaggctgagctcatggccctcagca 4219  
Db 2221 ggcagcagcctccggaaagaaacttcagcgcaaaaggctgagctcatggccctcagca 2280  
QY 4220 agcttggcgtggcggaaggaataccataaactttatagcagcaggtatgctt 4279  
Db 2281 agcttggcgtggcggaaggaataccataaactttatagcagcaggtatgctt 2340  
QY 4280 tgcgactgcacagtcacacgggccccttataaacaaggggtgttcttaccctcagcag 4339  
Db 2341 tgcgactgcacagtcacacgggccccttataaagaaggggtgttcttaccctcagcag 2400  
QY 4340 gagggaataaaagaaaggaataatttaagccttattagaagccttaccatttgcacaa 4399

Db 2401 gagggaaataaagaaagaaagaaattcttaagcctattagaagccttacatttgcacaa 2460  
 QY 4400 aaggttagctatatatacactgtcttgacatcagaaagccaaagatctcatatcagagg 4459  
 Db 2461 aaggttagctatatatacactgtcttgacatcagaaagccaaagatcccatatccagagg 2520  
 QY 4460 gaaccagatggtgaccgggttgccaaagcaggcagccaggctgttaacctctgcctat 4519  
 Db 2521 gaaccagatggtgaccgggttgccaaagcaggcagccagggtgttaacctctgcctat 2580  
 QY 4520 aatagaacccgcccagccagaaacccagacagacagtacaccctagaagactggcaaga 4579  
 Db 2581 gatagaacaccccagccagaccggcgacagacacccctagaagactggcaaga 2640  
 QY 4580 gataaaagatagaccagttctctgagactccggagggagcctgctatacctcatagg 4639  
 Db 2641 gataaaagatagaccagttctctgagactccggagggacctgctatacctcagatgg 2700  
 QY 4640 gaaggaatcctgcccacaaagaggttagaataatgtccaaacagatatacctgttaac 4699  
 Db 2701 gaaggaatcctgcccacaaagaggttagaataatgtccaaacagatatacctgttaac 2760  
 QY 4700 ccacttaggaactaaacactgcagcagtggtgcagaacatcccttatcatgttctgag 4759  
 Db 2761 ccacttaggaactaaacactgcagcagtggtgcagaacatctcttatcatgttctgag 2820  
 QY 4760 gctaccagagtggtgactcgtggtcacaacattgtgcccctgccagctgggttaatgc 4819  
 Db 2821 gctaccagagtggtgactcgtggtcacaacattgtgcccctgccagctgggttaatgc 2880  
 QY 4820 taactctccagaatacctccaggaagagactaagggaaggggaagccaccagcgctcactg 4879  
 Db 2881 taactctccagaatacctccaggaagagactaagggaaggggaagccaccagcgctcactg 2940  
 QY 4880 ggaagtgcactcactcaggttaaaagccggttaaatatcaggaacaaatctcttgggttt 4939  
 Db 2941 ggaagtgcactcactcaggttaaaagccggttaaatatcaggaacaaatctcttgggttt 3000  
 QY 4940 tctagacacccctttcaggtggttagaggttacttactactaaagaagagactcaaccgt 4999  
 Db 3001 tctagacacccctttcaggtggttagaggttacttactactaaagaagagactcaaccgt 3060  
 QY 5000 ggtggttaagaaatactggagaaatttttccaaagatttgaatacctaaggttaatagg 5059  
 Db 3061 ggtggttaagaaatactggagaaatttttccaaagatttgaatacctaaggttaatagg 3120  
 QY 5060 gtcagacaatggtccagcttctggtccaggttaagtcagggactggccaaagatttggg 5119  
 Db 3121 gtcagacaatggtccagcttctggtccaggttaagtcagggactggccaaagatttggg 3180  
 QY 5120 gattgattggaactgcatgtgcatacagaccaccaagctcaggacaggttagagaggtat 5179  
 Db 3181 gattgattggaactgcatgtgcatacagaccaccaagctcaggacaggttagagaggtat 3240  
 QY 5180 gaatagaaccattaaagaccccttacaattgacacagagactggcattaatgattg 5239  
 Db 3241 gaatagaaccattaaagaccccttacaattgacacagagactggcattaatgattg 3300  
 QY 5240 gatggctcctgccttctgctttttaggttgaggaaacccctggacaggttgggct 5299  
 Db 3301 gatgctcctgccttctgctttttaggttgaggaaacccctggacaggttgggct 3360  
 QY 5300 gacccctatgaattgctctacgggggacccccccctgtggcagaaattgcccctgcaca 5359  
 Db 3361 gacccctatgaattgctctacgggggacccccccctgttgtagaaattgcttcgtaca 3420  
 QY 5360 tagtgcctgctgcttctccagccttcttctcagctcaagcgctcagtggtggt 5419  
 Db 3421 tagtgcctgctgcttctccagccttcttctcagctcaagcgctcagtggtggt 3480  
 QY 5420 gaggcagcagctgggaagcagctccggagggcctactc---aggagagacttgcaggt 5476  
 Db 3481 gagacaacgagctgggaggaactccggagggcctactcagggagagagacttgcaggt 3540

QY 5477 tccacatcgtcttccaaagttggagattcagctctatgttagacgccacccgtgcagaaacct 5536  
 Db 3541 ccacatcgtttccaaagttggagattcagctctatgttagacgccacccgtgcagaaacct 3600  
 QY 5537 cgagactcgttggaaaggaccttatctctcgtactcttggaccacacccagcgtgtgaaagt 5596  
 Db 3601 cgagactcgttggaaaggaccttatctctcgtactcttggaccacacccagcgtgtgaaagt 3660  
 QY 5597 cgaaggaatccacacctggatccatgcacccagtttaagccggccacactccgactc 5656  
 Db 3661 cgaaggaatccacacctggatccatgcacccagtttaagccggccacactccgactc 3720  
 QY 5657 ggggtggagagccgaaagactgagaaaccccttaagcttcgctccactcgtcgtggtcc 5716  
 Db 3721 ggggtggagagccgaaagactgagaaaccccttaagcttcgctccactcgtcgtggtcc 3780  
 QY 5717 ttactctaaataaactcccccaggttagtaaacgccttatagacagctcgaaccccc 5776  
 Db 3781 ttactctaaataaactcccccaggttagtaaacgccttgtagacagcccgaaacctcc 3840  
 QY 5777 atagacctttatcccttaccctgctgattattgaacctgatacgggtgcactgtaataa 5836  
 Db 3841 ataaacctttatcccttaccctgctgattattgaacctgatacgggtgcactgtaataa 3900  
 QY 5837 gcaactgaggtgtgtcctcctagaggcacctggtggcctgaactgcatttctgcctccgat 5896  
 Db 3901 gcaactgaggtgtgtcctcctagaggcacctggtggcctgaactgcatttctgcctccgat 3960  
 QY 5897 tgattaaaccccgctgta-----aaagacacccctcccaacactagtcctgtagtgggt 5950  
 Db 3961 cagtaatccctggtctcaatgaccagggccacaccccgatgtactccgtgcttacgggt 4020  
 QY 5951 tctatgctgcccagg---cacagaaagagaataactgtgggttctctggggaactcct 6007  
 Db 4021 ttacggttggccagaccccccaataatgaagaaatattgtggaatacctcaggattct 4080  
 QY 6008 tctgtaggagatgagctgcctccaccccaacgattgagactggaaatggcgcgtctctc 6067  
 Db 4081 ttgcaagcaatgagctgcgttaactcttaattgagtggaattggaatggcgcgtctctc 4140  
 QY 6068 tcagagaccgggttaaaattctcttctgcaattcc----- 6102  
 Db 4141 agcaagacagagttaattctcttcttgaacaatcctaccagttataatcaatttaatt 4200  
 QY 6103 -----ggccggcgcaagtacaaatgatgaactataaa 6138  
 Db 4201 atggccatgggagatgaaagattggcaacagcgggtacaaaagatg-tacgaaataag 4259  
 QY 6139 gataagactgctcccccacagacttagattatctaaagataagtttccactg--aaagga 6196  
 Db 4260 caataagctgctcattcgttagacctagattacttaaaaaataagtttccactgaaaaagga 4319  
 QY 6197 aaacagaaaaataattcaaaagtgataaaatggtatgagctggggaatagtttttat 6253  
 Db 4320 aaacagaaaaataattcaaaagtggttaaatggtatgctgttggggaatagttactat 4376

## RESULT 14

AAV82749

ID AAV82749 standard; DNA; 3482 BP.

XX AAV82749;

XX 25-FEB-1999 (first entry)

DE Pig endogenous retrovirus (PERV)-B envelope (env) gene region.

XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
 KW probe; primer; detection; retrovirus; human tissue; xenotransplant;  
 KW primary porcine tissue; human cell line; porcine cell line; ss.

OS Pig endogenous retrovirus.

XX Key Location/Qualifiers  
 FT CDS 911..2884  
 FT /\*tag= a  
 FT /product= envelope\_protein

XX WO9853104-A2.

PD 26-NOV-1998.

XX 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

PI WPI: 1999-045324/04.

DR P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridising to either the PERV-A or PERV-B env gene - useful in the detection of retroviruses, and their subtypes, in a sample of porcine/human tissue

XX Claim 3; Page 21-23; 36pp; English.

CC The present sequence encodes a Pig endogenous retrovirus (PERV)-B envelope protein. PERV exists in two different subtypes, PERV-A and PERV-B. The differences are reflected in sequence divergence in the envelope genes. Probes and primers can be derived from the envelope (env) genes of PERV-A and PERV-B. The probes and primers are used in a method to detect retroviruses in a sample of porcine/human tissue, particularly primary porcine tissue and human cell lines that have been cultivated in the presence of a porcine cell line, or human tissue from a patient with a xenotransplant. Subtype of PERV in a sample containing one of the PERV env genes can also be determined.

XX Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

Query Match 41.1%; Score 3376.6; DB 20; Length 3482;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 3441; Conservative 0; Mismatches 29; Indels 7; Gaps 4;

QY 4718 cctgcagcagttgtgcagaaacatcccttcatcatgttctgagctaccagagtggtga 4777  
 |||||  
 DB 6 cctgcagcagttgtgcagaaacatcccttcatcatgttctgagctaccagagtggtga 65  
 QY 4778 ctgggtggtcaaacattgtgtccctgcagctggttaattgcttaactctccagaatacc 4837  
 |||||  
 DB 66 ctgggtggtcaaacattgtgtccctgcagctggttaattgcttaactctccagaatacc 125  
 QY 4838 tcagagaaagagactaagggaaggaacaccagcgcctcactgggaagtgacttcaactga 4897  
 |||||  
 DB 126 tccaggaagagactaagggaaggaacaccagcgcctcactgggaagtgacttcaactga 185  
 QY 4898 ggtaaacggcgttaaacacagaaacaaatctattgtttttgtagacaccttttcagg 4957  
 |||||  
 DB 186 ggtaaacggcgttaaacacagaaacaaatctattgtttttgtagacaccttttcagg 245  
 QY 4958 atgggtgaggttactactactaagaaagagacttcaacogtgggtggttaagaaatact 5017  
 |||||  
 DB 246 atgggtgaggttactactactaagaaagagacttcaacogtgggtggttaagaaatact 305  
 QY 5018 ggaagaaattttccaaagatttggaaatacctaaggaataaggtcagacaatgtccagc 5077  
 |||||  
 DB 306 ggaagaaattttccaaagatttggaaatacctaaggaataaggtcagacaatgtccagc 365  
 QY 5078 ttctgttgcaggaagtagcagagactggccagagattgtgggattgtattggaaactgca 5137  
 |||||  
 DB 366 ttctgttgcaggaagtagcagagactggccagagattgtgggattgtattggaaactgca 425

QY 5138 ttgtgcatacagaccccaaaagctcagagcaggtagagaggtgaatagaaccattaaaga 5197  
 |||||  
 DB 426 ttgtgcatacagaccccaaaagctcagagcaggtagagaggtgaatagaaccattaaaga 485  
 QY 5198 gaccttaccaaattgaccacagagactggcattaatgattggatggctctctcctgacct 5257  
 |||||  
 DB 486 gaccttaccaaattgaccacagagactggcattaatgattggatggctctctcctgacct 545  
 QY 5258 tgtgttttttaagggtgaggaacacccctggacagttttaggtctgacctccctatgaattgct 5317  
 |||||  
 DB 546 tgtgttttttaagggtgaggaacacccctggacagttttaggtctgacctccctatgaattgct 605  
 QY 5318 ctacgggggaccccccccgcttggcagaaattgcctttgcacatagctgctgctgct 5377  
 |||||  
 DB 606 ctacgggggaccccccccgcttggtagaaattgctctgtacatagctgctgctgct 665  
 QY 5378 ttccagcctttgttctctaggtcgaagcgcctgcagtggtggagcgagcagcgtgaa 5437  
 |||||  
 DB 666 ttccagcctttgttctctaggtcgaagcgcctgcagtggtggagcgaagcgtgaa 725  
 QY 5438 gacgtccgggagggcctactcagg--aggagacttgaagtccacatcgtctccaagt 5494  
 |||||  
 DB 726 gacgtccgggagggcctactcaggagagagacttgaagtccacatcgtctccaagt 785  
 QY 5495 tggagattcagctctatgttagacgccacgtgcaggaacacctgcagactcgtgtgaaagg 5554  
 |||||  
 DB 786 gggagattcagctctatgttagacgccacgtgcaggaacacctgcagactcgtgtgaaagg 845  
 QY 5555 acctatctcgtacttttgaccacacacacgctgtgaaagtgcgaaggaatccccacctg 5614  
 |||||  
 DB 846 ccttattctcgtacttttgaccacacacacgctgtgaaagtgcgaaggaatccccacctg 905  
 QY 5615 gatcatgcctccacgttaagccggccacacctcccagctcgggtgagagcgacga 5674  
 |||||  
 DB 906 gatcatgcctccacgttaagctggccacacctcccagctcgggtgagagcgacga 965  
 QY 5675 gactgagaatcccttaagcttcgctccatcgcctggttcttactactacaataaactc 5734  
 |||||  
 DB 966 gactgagaatcccttaagcttcgctccatcgcctggttcttactactacaataaactc 1025  
 QY 5735 cccagggcagtagtaaacgccttatagacagctcgaaccccccataagactttatccctta 5794  
 |||||  
 DB 1026 cccagggcagtagtaaacgccttatagacagctcgaaccccccataagactttatccctta 1085  
 QY 5795 cctgctgattattgacctgatacgggtgtcaactgtaataagactcagaggtgtgctc 5854  
 |||||  
 DB 1086 cctgctgattattgacctgatacgggtgtcaactgtaataagactcagaggtgtgctc 1145  
 QY 5855 ctgagggcacctgtgtggcctgaactgatttctgctccgattgattaaaccccgctgta 5914  
 |||||  
 DB 1146 ctgagggcacctgtgtggcctgaactgatttctgctccgattgattaaaccccgctgta 1205  
 QY 5915 aaagcacacacctcccaacctagtcctgattgggtttctattgtgtccagggcacagaga 5974  
 |||||  
 DB 1206 aaagcacacacctcccaacctagtcctgattgggtttctattgtgtccagggcacagaga 1265  
 QY 5975 aagagaaatctgtgtgggtttctgggaaatccttctgttaggagatggagctgctcaact 6034  
 |||||  
 DB 1266 aagagaaatctgtgtgggtttctgggaaatccttctgttaggagatggagctgctcaact 1325  
 QY 6035 ccaacgagtaggagactggaaatggcagctctctccagaccgggtgaaatctctcttgg 6094  
 |||||  
 DB 1326 ccaacgagtaggagactggaaatggcagctctctccagaccgggtgaaatctctcttgg 1385  
 QY 6095 tcaattccggccggcgaagtagtaaaatgatgaaactatataaagataagagctgctccc 6154  
 |||||  
 DB 1386 tcaattccggccggcgaagtagtaaaatgatgaaactatataaagataagagctgctccc 1445  
 QY 6155 catcagacttagattatctaaagataagtttcaactg--aaagaaaaaacaggaaatattc 6212  
 |||||  
 DB 1446 catcagacttagattatctaaagataagtttcaactgaaagaaagaaagaaatattc 1505



QY 6213 aaaagtgatgataaagtgtatgagctggggaatagttttttat-tatatgggggggagcag 6271  
DB 1506 aaaagtgatgataaagtgtatgagctggggaatagtttttttaaaatatgggggggagcag 1565  
QY 6272 ggtccactttaaccattcgcttagatagagacggggagacaaacccctgtgccaatgg 6331  
DB 1566 ggtccactttaaccattcgcttagatagagacggggagacaaacccctgtgccaatgg 1625  
QY 6332 gacccgataaagtagtggctgaacaggggccccgggccccctggagccaccgcgataaactgc 6391  
DB 1626 gacccgataaagtagtggctgaacaggggccccgggccccctggagccaccgcgataaactgc 1685  
QY 6392 cgggtgcccccaattaaacctcgctgagccctgacataaacacagcccgcttagcaacagtagtacca 6451  
DB 1686 cgggtgcccccaattaaacctcgctgagccctgacataaacacagcccgcttagcaacagtagtacca 1745  
QY 6452 ctgattgattctctaccacacgcttagaactccccaggtgttcctgtttaagacagagac 6511  
DB 1746 ctgattgattctctaccacacgcttagaactccccaggtgttcctgtttaagacagagac 1805  
QY 6512 agagactcttcagttctcatccaggaggtcttcccaagccatcaactccaccgacccctgatg 6571  
DB 1806 agagactcttcagttctcatccaggaggtcttcccaagccatcaactccaccgacccctgatg 1865  
QY 6572 ccaacttcttctgtgttctatctctcagggcctccttattatgagggatgcta 6631  
DB 1866 ccaacttcttctgtgttctatctcctcagggcctccttattatgagggatgcta 1925  
QY 6632 aagaaagaaaattcaattgtgacaaagagatgaaatacaatgtacatgggggtcccgaa 6691  
DB 1926 aagaaagaaaattcaattgtgacaaagagatgaaatacaatgtacatgggggtcccgaa 1985  
QY 6692 ataagttaccctcactgaagtttccgggaaggagacatgcatagggaaagctcccccat 6751  
DB 1986 ataagttaccctcactgaagtttccgggaaggagacatgcatagggaaagctcccccat 2045  
QY 6752 cccaccacaccttctgtatagtagtctgtttagagagggcctcagaaatacagatt 6811  
DB 2046 cccaccacaccttctgtatagtagtctgtttagagagggcctcagaaatacagatt 2105  
QY 6812 tagtacctgggtataacagggtgggggcatgcaatactgggttaaacccctgtgttcca 6871  
DB 2106 tagtacctgggtataacagggtgggggcatgcaatactgggttaaacccctgtgttcca 2165  
QY 6872 cctcagttctcaacaaatccaaagattctgtgtcattggttccaaatcgtcccccgagt 6931  
DB 2166 cctcagttctcaacaaatccaaagattctgtgtcattggttccaaatcgtcccccgagt 2225  
QY 6932 actaccatcctgaggaaggtgctctgataaataagtagtctgttagcaacccagaaaaa 6991  
DB 2226 actaccatcctgaggaaggtgctctgataaataagtagtctgttagcaacccagaaaaa 2285  
QY 6992 gagaacccgtatcccttaccctagctgtaattgctcgattaggaagccgctgtggcgtag 7051  
DB 2286 gagaacccgtatcccttaccctagctgtaattgctcgattaggaagccgctgtggcgtag 2345  
QY 7052 gaacaggagacgctgccttgatcacaggaaccacagcagctagagaaggacttggtagc 7111  
DB 2346 gaacaggagacgctgccttgatcacaggaaccacagcagctagagaaggacttggtagc 2405  
QY 7112 tacatcgccgcatgacagaagatctccagagcccttaaaaggaggtctgttagcaacccagaag 7171  
DB 2406 tacatcgccgcatgacagaagatctccagagcccttaaaaggaggtctgttagcaacccagaag 2465  
QY 7172 agtccctgacttcttctgtaagtggttctacagaaccggagggattagatctcgt 7231  
DB 2466 agtccctgacttcttctgtaagtggttctacagaaccggagggattagatctcgt 2525  
QY 7232 ttctaagagagggtgggttattgtcagcccttaaaagaagaattgttcttatgtagatc 7291  
DB 2526 ttctaagagagggtgggttattgtcagcccttaaaagaagaattgttcttatgtagatc 2585  
QY 7292 actcaggagccatcagagactccatgaacaaagcttagaaaaagtttagagagcgctcgaa 7351

DB 2586 actcaggagccatcagagactcactagcaagcttagaagaaggttagagagcgctcgaa 2645  
QY 7352 ggaagaagagagcgtgaccaggggtgttgaagatggttcaacaggtctccttgatga 7411  
DB 2646 ggaagaagagagcgtgaccaggggtgttgaagatggttcaacaggtctccttgatga 2705  
QY 7412 ccaccctgcttctgctgacggggccccctagtagtctcctgttacttaacagtgc 7471  
DB 2706 ccaccctgcttctgctgacggggacccctagtagtctcctgttacttaacagtgc 2765  
QY 7472 ggcttgccttaataataggtttgttgcctttgttagagaacgagtgagtcagtcocaga 7531  
DB 2766 ggcttgccttaataataggtttgttgcctttgttagagaacgagtgagtcagtcocaga 2825  
QY 7532 tcatgtactctaggcaaacagtagccaagggccttctgagccaaggagaaactgacctagc 7591  
DB 2826 tcatgtactctaggcaaacagtagccaagggccttctgagccaaggagaaactgacctagc 2885  
QY 7592 ctcccgagttcgaagattagaactattaacaagacaagaagtggggaaatgaaggatgaa 7651  
DB 2886 ctcccgagttcgaagattagaactattaacaagacaagaagtggggaaatgaaggatgaa 2945  
QY 7652 aatgcaacctaaacctcccagaccaggaagttaataaaagcttaaatgcccccgaa 7711  
DB 2946 aatgcaacctaaacctcccagaccaggaagttaataaaagcttaaatgcccccgaa 3005  
QY 7712 ttacagacctgctggtccagtagtaaataggtagaaggtcacacttctattgttccag 7771  
DB 3006 ttacagacctgctggtccagtagtaaataggtagaaggtcacacttctattgttccag 3065  
QY 7772 gctgctactcctggcctgaagtaacaggaatgagttgactaactatcgcttatctgga 7831  
DB 3066 gctgctactcctggcctgaagtaacaggaatgagttgactaactatcgcttatctgga 3125  
QY 7832 ttctgtaaactgactggaacacatagagaatgattacacattgacagccctagtgacc 7891  
DB 3126 ttctgtaaactgactggaacacatagagaatgattacacattgacagccctagtgacc 3185  
QY 7892 tatctcaactgcaatctgctcactctgccagagagccacagcagatgcggaccccgagc 7951  
DB 3186 tatctcaactgcaatctgctcactctgccagagagccacagcagatgcggaccccgagc 3245  
QY 7952 tattttaaatgattggtccacggagcgcggtcctcgatattttaaaatgattggtcca 8011  
DB 3246 tattttaaatgattggtccacggagcgcggtcctcgatattttaaaatgattggtcca 3305  
QY 8012 tggagcgcggtgc-tcgatattttaaatgattggtttgtgacgcacaggtttgtgt 8070  
DB 3306 cggagcgcggtgccttcgatatattttaaatgattggtttgtgacgcacaggtttgtgt 3365  
QY 8071 gaacccataaaagctgtcccgattccgactcggggcgccgagtccttaccctcgctg 8130  
DB 3366 gaacccataaaagctgtcccgattccgactcggggcgccgagtccttaccctcgctg 3425  
QY 8131 gtgtacgactgtggggcccccgagcggttgggaataaaatcctctgtgtttgcatc 8187  
DB 3426 gtgtacgactgtggggcccccgagcggttgggaataaaatcctctgtgtttgcatc 3482

## RESULT 15

AAC67022

ID AAC67022 standard; DNA; 4918 BP.

XX AAC67022;

XX AAC67022;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 22.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.



```
XX WO200071726-A1.
PN
XX
XX 30-NOV-2000.
PD
XX
XX 24-MAY-2000; 2000WO-US14296.
PF
XX
XX 24-MAY-1999; 99US-0135631.
PR
XX
XX (MAYO-) MAYO MEDICAL VENTURES.
PA
XX
XX Federspiel MJ;
PI
XX
XX WPI; 2001-032041/04.
DR
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells -
XX
XX Claim 16; Page 109-111; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient.
XX
XX Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;
SQ

Query Match 40.1%; Score 3295.2; DB 22; Length 4918;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 4044; Conservative 0; Mismatches 633; Indels 229; Gaps 14;

QY 3437 tggatttcagactgtggatcccggtttgcgaccttagcagccccacctaccogct 3496
DB 1 tggatttcagactgtggatcccggtttgcgaccttagcagccccacctaccogct 60
QY 3497 aacaaagaaaggggattctctggctctgagcaccagaaagccatttgatgctat 3556
DB 61 aacaaagaaaggggattctctggctctgagcaccagaaagccatttgatgctat 120
QY 3557 caaaaagccctgctgagcaccctgctgctgctccctgcgtaactaaaccccttac 3616
DB 121 caaaaagccctgctgagcaccctgctgctgctccctgcgtaactaaaccccttac 180
QY 3617 cctttatgtgagtcgagcgaagagtagcccgaggaggttttaacccaaacccctagacc 3676
DB 181 cctttatgtgagtcgagcgaagagtagcccgaggaggttttaacccaaacccctagacc 240
QY 3677 atggagagaccctgttgcctacctgctcaaaagagcttgcctgtagccagtggtggcc 3736
DB 241 atggagagaccctgttgcctacctgctcaaaagagctgctcctgtgagccagtggtggcc 300
QY 3737 cgtatgctgaagcgtatcgagcgtgtggccatactgtgcaagagcgtgacaaaattgac 3796
DB 301 cgtatgctgaagcgtatcgagcgtgtggccatactgtgcaagagcgtgacaaaattgac 360
QY 3797 ttgggacagaataactglaataagcccccatgcatgttggaagacatcttgcgagcc 3856
DB 361 ttgggacagaataactglaataagcccccatgcatgttggaagacatcttgcgagcc 420
QY 3857 ccagaccgagtgatgacaaaccccgcatgacccactatcaagcctgcttctcagaga 3916
DB 421 ccagaccgagtgatgacaaaccccgcatgacccactatcaagcctgcttctcagaga 480
QY 3917 gaggtcacttgcctccacagccgctctcaacccctgccactcttctgctgagagagac 3976
DB 481 gaggtcacttgcctccacagccgctctcaacccctgccactcttctgctgagagagac 540
QY 3977 tgatgaaccagtgactcatgttgccatcaactattgattgagagactgggtccgcga 4036
DB 1 tgatgaaccagtgactcatgttgccatcaactattgattgagagactgggtccgcga 1680

DB 541 tgatgaaccagtgactcatgttgccatcaactattgattgagagactgggtccgcga 600
QY 4037 ggaccttacagacataccgctgacigggagaagtgtcaacctgttcaactgacggagcag 4096
DB 601 ggaccttacagacataccgctgacigggagaagtgttaacctgttcaactgacggagcag 660
QY 4097 ctatgtgtggaaggttaagaggtggtggcgagtggtggcgagtggtggcgagccgacacat 4156
DB 661 ctatgtgtggaaggttaagaggtggtggcgagtggtggcgagtggtggcgagccgacacat 720
QY 4157 ctggggccagcagccctgcccgaagaaacttcagcgcaaaaggtgagctcatggccctcac 4216
DB 721 ctggggccagcagccctgcccgaagaaacttcagcgcaaaaggtgagctcatggccctcac 780
QY 4217 gcaagcttgcggtggtgcccgaagaaacttcataaacattattacgagcagcagtgatgc 4276
DB 781 gcaagcttgcggtggtgcccgaagaaacttcataaacattattacgagcagcagtgatgc 840
QY 4277 ctttgcgactgcacacgtacacggggccatctataaaacaaaggggtgttgcctacacgc 4336
DB 841 ctttgcgactgcacacgtacacggggccatctataaaacaaaggggtgttgcctacacgc 900
QY 4337 agggagggaaataaagacaacaaagagaaattctaagcctattagaagcctacatttgc 4396
DB 901 agggagggaaataaagacaacaaagagaaattctaagcctattagaagcctacatttgc 960
QY 4397 aaaaaggtcagctattatatacactgtcctggacatcagaaaagcgaagatctcatatcag 4456
DB 961 aaaaaggtcagctattatatacactgtcctggacatcagaaaagcgaagatctcatatcag 1020
QY 4457 agggaaaccagatggctgaccgggttgccaagcagcagcccgaggtgttaaccttctgc 4516
DB 1021 agggaaaccagatggctgaccgggttgccaagcagcagcccgaggtgttaaccttctgc 1080
QY 4517 tataatagaacgcgcccacaaagcccccagacccagacgacgacgtacacccctagaagactgca 4576
DB 1081 tatgatagaacacccacaaagcccccagacccagacgacgtacacccctagaagactgca 1140
QY 4577 agagataaaaagatagaccagttctctgagactccgaggggagcctgtatcacctcata 4636
DB 1141 agagaataaaaagatagaccagttctctgagactccgagggagcctgtatcacctcata 1200
QY 4637 tgggaaggaatactctgcccacaaagaggttagaataatgtcccaacagatcacctgct 4696
DB 1201 tgggaaggaatactctgcccacaaagaggttagaataatgtcccaacagatcacctgct 1260
QY 4697 aacccacttagaactaaacacccctgagcagttggtcagaacatcccccttatctgtct 4756
DB 1261 aacccacttagaactaaacacccctgagcagttggtcagaacatcccccttatctgtct 1320
QY 4757 gaggtaccagagtggtgactcggtgtcaaacatgtgtgcccctgcccagctggttaa 4816
DB 1321 gaggtaccagagtggtgactcggtgtcaaacatgtgtgcccctgcccagctggttaa 1380
QY 4817 tgcataccttccagaataacccctccaggaagagactaaggggaagcccccagggcgctca 4876
DB 1381 tgcataccttccagaataacccctccaggaagagactaaggggaagcccccagggcgctca 1440
QY 4877 ctgggaagtggacttcaactgaggtaaagcccgctaaaatacgggaacaaatactattgtt 4936
DB 1441 ctgggaagtggacttcaactgaggtaaagcccgctaaaatacgggaacaaatactattgtt 1500
QY 4937 tttttagacacctttcagagtggtgaggtctatcctactaagaagaagacactcaac 4996
DB 1501 tttttagacacctttcagagtggtgaggtctatcctactaagaagaagacactcaac 1560
QY 4997 cgtgtgtgctagaataactgaggaataattttccaagatttgaataccttaaggtta 5056
DB 1561 cgtgtgtgctagaataactgaggaataattttccaagatttgaataccttaaggtta 1620
QY 5057 agggctcagacaatgttccagcttctggtccaggttaagtcaggagactggcgaagatt 5116
DB 1621 agggctcagacaatgttccagcttctggtccaggttaagtcaggagactggcgaagatt 1680
```



Search completed: February 24, 2002, 03:02:54  
 Job time: 19340 sec

```

Db 3789 agaaaggggttagatcgttatttctaaagaagagggttatgtgtagccttgaaggag 3848
QY 7270 gaatgtgcttcatatagatcctcagagccatcagagactcactgaacaagcttaga 7329
Db 3849 gaatgcgttttattgtatcattcaaggccatcagagactcactgaacaagcttaga 3908
QY 7330 aaaaagttagagagcgctcgaaggaagaagagcgctcgaagaggtgtgttgaagatgg 7389
Db 3909 gaaaggttgagagagcgctcgaaggaagaagaaactcactcaaggtgtgttgaagagatgg 3968
QY 7390 ttcaaaaggtctccttgatgataccacgtccttctcgtctgaagcgggcccttagatgc 7449
Db 3969 ttcaaaaggtctccttgatgataccacgtccttctcgtctgaagcgggcccttagatgc 4028
QY 7450 ctgctccttgatcactgaatggcgcttgcttaataatgaagttgtgtccttgttaga 7509
Db 4029 ctccctccttgatcactgaatggcgcttgcttaataatgaagttgtgtccttgttaga 4088
QY 7510 gaacgagtgatgacgtccagatcagtgacttaggcaacagttaccaagcgcttcagc 7569
Db 4089 gaacgagtgatgacgtccagatcagtgacttaggcaacagttaccaagcgcttcagc 4148
QY 7570 caagagaaactgacccctcagcctccagttcctaagattgaactatlaacaagacaag 7629
Db 4149 ---agagaagcgcgctgactcactcagttcctaagattgaactatlaacaagacaag 4205
QY 7630 aagtgaggatgaagaatgaagaatgaacccctccagaaaccaggaagttata 7689
Db 4206 aagtgaggatgaagaatgaagaatgaacccctccagaaaccaggaagttata 4265
QY 7690 aaaaagctcctaaatgccccgaattacagacccctgctgcgcagtaaatagtgtagaag 7749
Db 4266 agaaagctcctaaatgccccgaattacagacccctgctgcgcagtaaatagtgtagaag 4317
QY 7750 tcaacattcctattgttccagggcctgctatccctgagcctaagtaagtaacaagaaatga 7809
Db 4318 tcaacattcctattgttccagggcctgctatccctgagcctaagtaagtaacaagaaatga 4377
QY 7810 gttaga-----ctaatc 7820
Db 4378 aaaaacagggccctgagtgatgtgctctatgtcttgaacactcttgaacactgctcctaact 4437
QY 7821 gcttatcgtgattctgttaaacctgactg----- 7848
Db 4438 gcttgcttgcttctgttaaacctgcttgcataagataaaagagagagatgcaattgctt 4497
QY 7849 ---gcaccatagaagaatgtgattacatgtgacagccctagtgacctatctcaactgcaa 7905
Db 4498 aacggaaccccgtaagatcggtgtgtacacaacaatgttgaaacacatatcttggtagaaa 4557
QY 7906 tctgtcactcgcgcagagcccaacgacatgcgagacctccgagcctatttlaaatgat 7965
Db 4558 catgtctcccccaccccgaaacatgcgcaaatgtgtactcttaaacatatttaataat 4617
QY 7966 tggtcacagagagcgcggtctcgtatatttlaaaaatgattgtgtccatggaagcgcggtc 8025
Db 4618 tggtcacagagagcgcggtctcgtatatttlaaaaatgattgtgtccatggaagcgcggtc 8085
QY 8026 tcatatttlaaaatgattgtgtgtgacacagcgcttgtgtgaaccccatataaagc 8085
Db 4665 --gatatttgaatgtattgtgtgtgaagcgcggttgtgtgaaccccatataaagc 4722
QY 8086 tgtccgattccgactcgcggcgccgacgtccttaacccctgcgtgtgtgaaccccatataaagc 4722
Db 4723 tgtccgactcgcacactcgcggcgccgacgtccttaacccctgcgtgtgtgaaccccatataaagc 4782
QY 8146 ccccaagcgcgcttggaataaaatcctctgctgtgtgtgcatcaaaa 8191
Db 4783 ccccaagcgcgcttggaataaaatcctctgctgtgtgtgcatcaaaa 4828

```





```

FT      protein_bind      7885..7914
FT      /*tag= p
FT      /note= "Binding site for transcription factors
FT      AP-1/GATA, c-Myc and AP-1"
FT      protein_bind      7928..7953
FT      /*tag= q
FT      /note= "Binding site for transcription factors E47
FT      and ETS-1"
FT      CAAT_signal      7964..7968
FT      /*tag= r
FT      /note= "Binding site for transcription factor GATA"
FT      protein_bind      7985..7997
FT      /*tag= s
FT      /note= "Binding site for transcription factor GATA"
FT      CAAT_signal      8003..8007
FT      /*tag= t
FT      protein_bind      8024..8036
FT      /*tag= u
FT      /note= "Binding site for transcription factor GATA"
FT      CAAT_signal      8042..8046
FT      /*tag= v
FT      protein_bind      8049..8060
FT      /*tag= w
FT      /note= "Binding site for transcription factor
FT      AP-1/CREB"
FT      TATA_signal      8075..8090
FT      /*tag= x
FT      misc_feature      1..61
FT      /*tag= y
FT      /note= "R LTR domain"
FT      misc_feature      8127..8135
FT      /*tag= z
FT      /note= "Polya downstream element"
FT      polya_signal      8161..8166
FT      /*tag= aa
FT      W09740167-A1.
FT      30-OCT-1997.
FT      PD
FT      18-APR-1997; 97WC-GB01087.
FT      PF
FT      10-FEB-1997; 97GB-0002668.
FT      PR
FT      19-APR-1996; 96GB-0008164.
FT      XX
FT      PA
FT      (IMUT-) IMUTRAN LTD.
FT      (QONE-) Q-ONE BIOTECH LTD.
FT      XX
FT      PI
FT      Galbraith DN, Haworth C, Lees GM, Smith KT;
FT      WPI: 1997-535851/49.
FT      DR
FT      P-PSDB; AAM39271; AAM39272; AAM39273.
FT      XX
FT      PT
FT      Polynucleotide encoding porcine retrovirus expression product -
FT      useful to develop products for use in vaccines, diagnosis and
FT      xeno-transplantation
FT      PT
FT      Claim 4; Fig 3; 69pp; English.
FT      PS
FT      XX
FT      CC
FT      This DNA sequence encodes the porcine retrovirus (PoRV) virion core
FT      CC
FT      polypeptide (GAG), polymerase (POL), and envelope (ENV) proteins and
FT      CC
FT      also includes the Long Terminal Repeat (LTR). These proteins can be used
FT      CC
FT      to develop viral vaccines, antisense nucleic acids, ribozymes and other
FT      CC
FT      antiviral agents. They can also be used in xeno-transplantation
FT      CC
FT      technology and as diagnostic tools.
FT      XX
FT      SQ
FT      Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

```

alignment\_scores:  
 Quality: 2801.00 Length: 524  
 Ratio: 5.345 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-171-553B-4 x AAV09700 ..  
 Align seg 1/1 to: AAV09700 from: 1 to: 8209

```

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17
|||||
588 ATGGGACAGACAGTACTACCCCCCTTAGTTGACCTGACCATTTGGAC 637
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
638 TGAAGTTAGATCCAGGCTCATATTGTCAGTTGACGTTTACAGAGGAC 687
34 rGTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
688 CTGGGACAGACTTCTGTGCTGCTGAATGGCCACATTCGATGTTGATG 737
51 ProSerGluGlyThrPheAsnSerGluLeuAlaValLysAla11 67
|||||
738 CCATCAGAGGGGACCTTAATTCGAATATTCCTGCTGTTAAGCAAT 787
67 eLlePheGlnThrGlyProGlySerHisProAspGlnGluProTrpLle 84
|||||
788 CATTTTCAGACTGAGCCGCGCTCATCTGATCAGAGCCCTATATTC 837
84 eutHrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
|||||
838 TTACGTGGCAAGATTGGCAGAGATCCCGCATGGGTTAAACCATG 887
101 LeuAsnLysProArgLysProGlyProArgLleLeuAlaLeuGlyLys 117
|||||
888 CTAAATTAACCAAGAAAGCCAGGTCCTCCCAATCTGCTTGGAGAGAA 937
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyLeuPro 134
|||||
938 AACCAAACTCGGCCGCAAAAAGTCAGCCCTTCTCTGATCTACCCC 987
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
988 GAGATCGAGAGCGCGCGACTGTGGCGGAGACCCCAACCTGTCGCCACC 1037
988 GAGATCGAGAGCGCGCGACTGTGGCGGAGACCCCAACCTGTCGCCACC 1037
151 ProLeuSerSerThrGlyCysGlyGlyThrSerAlaProProGlyAla 167
|||||
1038 CCTTATCCAGACAGGCTGCTGAGGGGACCTGCCCCCTCTGAGAG 1087
167 aProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla17 184
|||||
1088 TCCGGTGTGGAGGAGACTCTGCGCGAGACTCGGAGCGGAGAGCGCCA 1137
184 hrProGluArgThrAspGluLleAlaLleLeuProLeuArgThrTyGly 200
|||||
1138 CCGCGAGCGGACAGACAGAGATCGCATATTACCGCTGCGCACCTATGCG 1187
201 ProPrometProGlyGlnLeuGlnProLeuGlnTyTrpProPhe 217
|||||
1188 CCTCCATGCGAGGGGCGCAATTGCGACCCCTCAGATTGCGCCCTTTTC 1237
217 rSerAlaAspLeuTyAsnTrpLysThrAsnHisProProPheSerGlu 234
|||||
1238 TTTCGCGAGATCTCATATTGGAATACTACATCCCCCTTCTCGGAGG 1287
234 sPProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGln 250
|||||
1288 ATCCCAAGCGCTCAGCGGGGTGGTGAGTCCCTTATGTTCTCTCAGCAG 1337
251 ProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLéuPheThrArg 267
|||||
1338 CCTACTTGGGATGATTGTCAACAGCTGCTCAGACACTCTTCAACACCG 1387
267 uGluArgGluArgLleLeuLeuGluAlaArgLysAsnValProGlyAla 284
|||||
1388 GGACGAGAGAGAAATTTCTTTAGAGGCTAGAAAAAATGTTCTCTGGGCG 1437

```

```

284 spgIyArpProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLeu 300
1438 ACGGGCAGCCACGAGTTCGAAATGATTCATGGGATTCCTTG 1487
301 ThrArgProGlyTTPAspTyrAsnThrAlaGluGlyArgGluSerLeu 317
1488 ACTGGCCCCGGGTGGAGCTACACACGCTGAAAGTAGAGGAGAGCTTGA 1537
317 sILeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgp 334
1538 AATCATGCCAGGCTCTGTGGGGGTCTCGGGGGCCCTCAAGACGAGC 1587
334 rorThAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPro 350
1588 CCACTAATTTGGCTAGGTAGAGAGGTGATCGAGGCGCAAGAACCT 1637
351 ProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
1638 CCTCGGTATTTCTTGAGAGGCTCATGGAAGCTTCAGGGGCTCACCC 1687
367 oPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 384
1688 TTTGATCTTACTCTAGAGGCCCGCAAGACCTCAGTGGCTGGCTTCA 1737
384 IeGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGly 400
1738 TTGGGCACATGGCTCTGTGATTCAGAAACCTTCAGAGAGCTGGAAGG 1787
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValTyr 417
1788 TTACAGGAGGCTGAGTTACGTGATCTAGTGAGAGAGGCGAGAGGTCT 1837
417 rTyArgArgGluThrGlnGluGlyLysGlnArgLysLysLysGlu 434
1838 TTACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1887
434 rGgGluArgGluGluArgArgAspArgGlnGluLysAsnLeuThr 450
1888 GAGAAAGAGGAGGAGAAAGAGCTGATAGCGCAAGAGAAAGAAATTGACT 1937
451 LysIleLeuAlaIleValIleGlyLysSerSerArgGluArgGluArg 467
1938 AAGATCTTGCCCGCAGTGTGAAGGAGAGCAGAGGAGAGAGAGAG 1987
467 gAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 484
1988 AGATTTTAGGAAATTAGGTAGGCTCCTAGACAGCTCAGGGAGCTGGG 2037
484 snATGThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
2038 ATAGAGCCCGCAGTGCAGACAGGACAGTGTGCTATTGTAAGAAAGCA 2087
501 HisThrAlaArgAsnGlyProLysLysGlyAsnLysGlyProLysVal 517
2088 CACTGGGCAAGCACTGCCCAAGAGAGGAAAGAGGAGGAGGAGGCT 2137
517 uAlaLeuGlnGluAspLysAsp 524
2138 AGCTCTAGAGAGAGATTAAGAT 2159
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699
seq_documentation_block:
ID AAV09699 standard; DNA; 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

```

```

XX Porcine retrovirus.
OS
XX
FH Key Location/Qualifiers
FT CDS 576..2126
FT /tag= a
FT /product= GAG protein
FT /note= "viroin core polypeptide"
FT CDS 2143..5733
FT /tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
specification"
FT CDS 506..7576
FT /tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX
PD WO9740167-A1.
XX
PF 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMOT-) IMOTRAN LTD.
PA (QONE-) Q-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees CM, Smith KT;
XX
DR WPI: 1997-535851/49.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation
XX
PS Claim 4; Fig 2; 69pp; English.
XX
CC This DNA sequence encodes the porcine retrovirus (POEV) viroin core
CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These
CC proteins can be used to develop viral vaccines, antisense nucleic acids,
CC ribozymes and other antiviral agents. They can also be used in
CC xeno-transplantation technology and as diagnostic tools.
XX
SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other:
alignment_scores:
Quality: 2774.00 Length: 524
Ratio: 5.304 Gaps: 0
Percent Similarity: 99.809 Percent Identity: 99.809
alignment_block:
US-09-171-553B-4 x AAV09699 ..
Align seg 1/1 to: AAV09699 from: 1 to: 8196
1 MetGlnGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
|||||
576 ATGGACAGACAGTACTACCCCTTACTGTTGACTCTGACCATTTGGAC 625
|||||
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
626 TGAAGTTAGATCCAGGCTCATTAATTGTTCAGTTCAAGTTAAGAGGAG 675
|||||
34 rorThGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
676 CTGGGACAGCTTCTGTGCTCTGAAATGGCCACATTCATGTTGATGG 725
|||||
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
|||||

```

```

726 CCATCAGAGGGGACCTTTAATTCGAAATTAATCTGCGCTGTTAAGCAAT 775
67 eilleheglnthrglyproglyserhisproaspglnluProTyrllel 84
|||||
776 CATTTTTCAGACTGAGACCGGCTTCATCTGATCAGAGAGCCCTATATCC 825
84 euThrTfpglnaspleuAagiuaaspproporproTfvallyProthr 100
|||||
826 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCAAGG 875
101 leuAenlyProarglysProglyProarglyleuAalaleuAglly 117
|||||
876 CTAATTAACCAAGAAAGCCAGAGTCCCGCAATCCGCGCTCTTGAGAGAA 925
117 sasnlyshisSerAagiulysValgluProSerSerTyrlleuProA 134
|||||
926 AAACAAACACTCGCGCAAAAGTCAGACCTCTCTCGTATCTACCC 975
134 rgaAspArglyAlaAlaAspleuAagiulYthrProthrCysSerProthr 150
|||||
976 GAGATCGAGAGAGCGCGCACTTGGCGGAGACCCCAACCTGTTCCCCACC 1025
151 ProleuSerSerThrglyCysCysgluYthrSerAlaProProglyAl 167
|||||
1026 CCTTATCCAGACAGGGGTGCTGTGAGGGAGACTCTGCCCTCTCGAGC 1075
167 aProValValgluYlProAlaAgiYthrArgSerArgArglyAlat 184
|||||
1076 TCCGCTGTGGAGGACCTGCTCCGGGAGTCGGAGCGGAGAGGCGCA 1125
184 hrProgluArgThraspgluIleAlaIleleuProleuArgThrTygly 200
|||||
1126 CCCCAGACGAGACAGAGATCGGATTAACGCTGCGCACTATAGGC 1175
201 ProProMetProglygluYlgluYlgluYlProleuYlYlYlYlYlYl 217
|||||
1176 CCTCCAGACGAGGGGCAATTCAGACCCCTCCAGATATGGCCCTTTC 1225
217 rSerAlaAspleuTyrrasntPlysrThrasnHisProProheSerGlu 234
|||||
1226 TTTCGCAATCTCTAATATGTGAAACTAACCAATCCCTTTCGAGG 1275
234 sPProgluArgleuThrglyleuValgluSerleuMetPheSerHisGln 250
|||||
1276 ATCCCAAGCGCTCACGGGGTGTGAGTCCCTTATGTCTCCACAG 1325
251 ProThrTrpasPasPcysGlnGlnleuYlgluYlYlYlYlYlYl 267
|||||
1326 CCTACTTGGGATGTGTCAACACCTGCTGCAGACACTCTTCAACAG 1375
267 uGluArgGluArgIleleuYlgluAatrglyAsnValProglyAla 284
|||||
1376 GGAGCGAGAGAAATCTGTAGAGGCTAGAAAAAATGCTCTGGGCGG 1425
284 spgLyArgProThrglnleuGlnasnGluIleAspMetGlyPheProleu 300
|||||
1426 ACGGCGACCCACCACTTGCATAATGATGACATGGGATTCCTTG 1475
301 ThrArgProglylYrPasPtyrAsnThralaGluYlYlYlYlYlYl 317
|||||
1476 ACTGCGCCCGGTGGACTACAAACAGCGTGAAGGTAGGAGAGCTTGA 1525
317 sIleTyArgGlnAlaIleuValAagiYlYlYlYlYlYlYlYlYl 334
|||||
1526 AATCTATCGCAGGCTGTGTGGGGGTCTCCGGGGCGCTCAAGAGCG 1575
334 roThrasnleuAlaIysValArggluValMetGlnGlyProAsnGluPro 350
|||||
1576 CCACTAATTTGGCTAAGGTAAAGAGGTGATGCAAGGAGCAAGAACT 1625
351 ProSerValPheleuGluArgleuMetGluAlaPheArgArgPheThr 367
|||||
1626 CCTCGGTAATTTTGTGAGAGGCTCATGAAAGCTTCAAGGGGTTCACCC 1675

```

```

367 oPheAspProthrSerGluAlaGlnIysAlaSerValAlaIleuAlaPheI 384
|||||
1676 TTTTGATCTTAACCTCAGAGGCCAGAAAGCCTCAGTGGCCCTGCTTCA 1725
384 leGlyGlnSerAlaIleuAspIleArgIysIysleuGlnArgleuGluGly 400
|||||
1726 TTGGGAGTGGGCTGTGATATCAGGAAGAACTTCAGAGACTGGAAGG 1775
401 leuGlnGluAlaIleuAargAspleuValArgGluAlaGluYlYlYl 417
|||||
1776 TTACAGAGGCTCAGATTACTGATCTTAGTACAGAGCGAGAGAGGTGTA 1825
417 rTyArgArgGluThrglnGlnGluYlYlYlYlYlYlYlYlYlYlYl 434
|||||
1826 TTACAGAGGAGAGACAGAAAGAGAGAGAAAGCAAGAAAGAAAGAGAG 1875
434 rglGluGluArgGluArgArgAspArgArgGlnGlnIlysasnleuThr 450
|||||
1876 GAGAAAGAAAGGAGAGAAAGCTGATAGACGGCAAGAAATTTTGACT 1925
451 LysIleuAlaAlaValAlaGluYlYlYlYlYlYlYlYlYlYlYlYl 467
|||||
1926 AAGATCTTGGCCCGCAGTGTGTAAGGAGAGAGCAGCAGGAGAGAGAG 1975
467 gaSpPheArgLysIleArgSerGlyProArgGlnSerGlyasnleuGly 484
|||||
1976 AGATTTAGAAATTAAGTCAAGCCCTAGACAGTCAGAGAACTTGCGCA 2025
484 snArgThrProleuAspIysAspGlnCysAlaIleTyrcysIysGluYl 500
|||||
2026 ATAGAGCCCTCCTCGACAAAGGACAGTGTGCTATTGTAAAGAAAAAG 2075
501 HisTrpAlaArgasnCysProIysIysGlyasnIysGlyProIysVal 517
|||||
2076 CACTGGGCAAGAACTGCCCAAGAAAGGAAACAAAGAACCCAGATC 2124
517 uAlaIleuGluGluAspLysasp 524
|||||
2125 ACCTCTAGAAAGAAAGATTAAGAT 2146

seq_name: /SID2/gcdata/geneseq/geneseqn/NA2001.DAT:AAF77726
seq_documentation_block:
ID AAF77726 standard; DNA; 7333 BP.
XX
AC AAF77726;
XX
DT 23-MAY-2001 (first entry)
XX
DE Defective retroviral genome isolated from PK-15 cell line.
XX
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
XX
SS.
XX
Unidentified.
XX
PN US6190861-B1.
XX
PD 20-FEB-2001.
XX
PP 13-DEC-1996; 96US-0766528.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
WP1: 2001-256211/26.
XX
DR P-PSDB: AAB73282, AAB73283, AAB73284.
XX
Assessing risk of endogenous retroviruses in clinical practice and in

```



PT xenotransplantation, comprises using probe sequences derived from swine  
 PT or miniature swine retroviral genome -  
 XX  
 PS Claim 1; Fig 2; 127pp; English.  
 XX

CC The present invention relates to a method for screening a cell or tissue  
 CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g. the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.

XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

# alignment\_scores:

Quality: 2758.00 Length: 525  
 Ratio: 5.263 Gaps: 1  
 Percent Similarity: 99.810 Percent Identity: 99.619

## alignment\_block:

US-09-171-553B-4 x AAF77726 ..

Align seg 1/1 to: AAF77726 from: 1 to: 7333

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrPth 17  
 598 ATGGGACAGACAGTACTACCCCTTAGTTGACTCTGACCATTCGAC 647  
 17 rgluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGly 34  
 648 TGAAGTTAGATCCAGGGCTCATTAATTTGTCACTTCAAGTTAAGAGGGAC 697  
 34 rOTrPdlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
 698 CTTGGCAGACTTCTGTGCTCTGAATGGCCACATTCGATTTGGATGG 747  
 51 ProSerGluGlyThrPheAsnSerLuiLeuAlaValLysAla 67  
 748 CCATCAGAGGGGACCTTAATTCGAAATTAATCCTGCTGCTTAAGCAAT 797  
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
 798 CATTTTCAGACTGAGCCGGCTCTCATCTGATCAGAGGCCCTATATCC 847  
 84 euThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100  
 848 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCATGG 897  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117  
 898 CTAAATTAACCAAGAAAGCCAGAGTCCCGCAATCTGCTTGAGAGAA 947  
 117 sasnLysHisSerAlaGluLysValGlnProSerSerSerLysLeuPro 134  
 948 AAACAACACCTCGCCGAAAGAGTGAACCCCTC.TCCCTGATCTACACCC 996  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 997 GAGATCGAGAGACCCCGACCTTGGCCGGAACCCCACTGTTCCCCACCC 1046  
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167  
 1047 CCTTATCCAGACAGGGTGTGTGAGGGGACCCCTGCCCCCTCTGAG 1096  
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183  
 1097 CTCGGTGTGAGGGACCTGCTGCCGGGACTCGAGCCGGAAGAGCGGCC 1146

184 ThrProGluArgPheAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
 1147 ACCCGGAGCGGACAGACAGATCGCGATATTACCGCTGCCACTATGG 1196  
 200 yProPoweTrProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217  
 1197 CCTTCCCATTCGACAGGGGCCAATTGACGCCCTCCAGATATGGGCCCTTT 1246  
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233  
 1247 CTTTCGAGATCTCTATTAATGAAACATAACCAATCCCTTCTCCGAG 1296  
 234 AspProGlnArgLeuThrGlyLeuValGlnSerLeuMetPheSerHisG 250  
 1297 GATCCCCACACCCCTCAGCGGGTGTGTGAGTCCCTTATGTTCTTCACA 1346  
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrTh 267  
 1347 GCCTACTTGGGATGATTGTCAACAGCTGCTGCACAGACTCTTCACACCG 1396  
 267 lgluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283  
 1397 AGAGCGAGAGAGAAATCTGTAGAGGCTAAAAAAATGTTCCCTGGGGCC 1446  
 284 AspGlyArgProThrGlnLeuGlnAsnGlnIleAspMetGlyPhePro 300  
 1447 GACGGCGCACCCACGACCTTCCAAATGACATTACATCGGATTTCCCTT 1496  
 300 uThrArgProGlyTyrTrpAspTyrAsnThrAlaGluGlyArgLysSerLeu 317  
 1497 GACTGCGCCCGGTGGGACTACACAGCGCTGAAGAGTGAAGAGAGTTGA 1546  
 317 yslIeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 1547 AAATCTATCGGCAGGCTCTGGTGGGGGCTCTCCGGCGGCTCAAGACGG 1596  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnLysProAsnGluPr 350  
 1597 CCCCACTAATTTGGCTAAGTAGAGAGTGAATGAGAGGACCGAAGCAAC 1646  
 350 oProSerValPheLeuGlnuArgLeuMetGluAlaPheArgArgPheThr 367  
 1647 TCCCTCGATTTCTTGAGAGGCTCATGAAAGCTTCAGCGGTTACACC 1696  
 367 ropPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 1697 CTTTGTATCTTACCTCAGAGGCCCGAAGACCTCAGTGGCCCTGGCTTC 1746  
 384 illeGlyInserAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400  
 1747 ATTGGGAGTCGGCTCTGATATTCAGAGAAACCTTGAGAGCTGGAGG 1796  
 400 yleuGlnGluAlaGluLeuValArgAspLeuValArgGluAlaGlyVal 417  
 1797 GTTACAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1846  
 417 yTYTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 433  
 1847 ATTACAGAAAGGAGACAGAAAGAGAGAAAGAGAAAGAGAAAGAGAG 1896  
 434 ArgGluGluArgGluGluArgGluArgGluArgGluGluLysAsnLeuTh 450  
 1897 AGAGAAAGAAAGGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946  
 450 rLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGlu 467  
 1947 TAAGATCTTGCCGCACTGTTGATGAGGAGAGAGAGAGAGAGAGAGAG 1996  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 1997 GAGATTTTACGAAATATAGTTCAGGCCCTAGACAGTACAGGAGAACCTG 2046

```

484 AsnArGThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500
|||||
2047 AATGAGACCCCACTCGACAAAGACAGCTGCTGATTTGTAAGAAAAAG 2096
500 YHISTrPAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal1 517
|||||
2097 ACACGCGGCAAGAGAACTGCCCAAGAAAGGAAACAAAGACCGAAGCTCC 2146
517 euAlaLeuGluGluAspLysAsp 524
|||||
2147 TAGCTCTAGAAAGATTAAGAT 2169

seq_name: /SID25/9/cgdata/geneseq/geneseq/NA1997.DAT:AA174883
seq_documentation_block:
ID AAT74883 standard; cDNA; 7393 BP.
XX
AC AAT74883;
XX
DT 09-FEB-1998 (first entry)
XX
DE Porcine retrovirus cDNA (defective).
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT /tag= b
FT /note= "putative GAG protein"
FT CDS 2320..4737
FT /tag= c
FT /note= "putative POL coding region (partial) as
FT /note= "described in the specification"
FT mat_peptide 2320..3522
FT /tag= d
FT /note= "putative POL protein (partial)"
FT mat_peptide 3516..4328
FT /tag= e
FT /note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /tag= f
FT CDS 4738..6725
FT /tag= g
FT /note= "putative ENV coding region (partial) as
FT /note= "described in the specification"
FT mat_peptide 4752..6722
FT /tag= h
FT /note= "ENV protein (partial)"
XX
PN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96MO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI, 1997-332804/30.
DR P-PDB; AAW32091-W32095.
XX
PT New nucleic acid from porcine retroviruses - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient

```

```

XX
PS Claim 16; Fig 2; 128bp; English.
XX
CC This cDNA sequence represents a defective purified swine retrovirus
CC found in PK-15 cells containing the putative coding regions for viral
CC GAG, POL and ENV proteins. There are a few in frame stop codons and
CC apparent frame shifts in the given coding sequence which alter features
CC of the translation. This sequence and PCR fragments generated from the
CC sequence (see AAT74812-T74882) could be used to screen organs for the
CC presence of porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.
CC copy number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and to
CC determine the histological localisation of activated retroviruses. Using
CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear
CC cells, infectivity titration and susceptibility testing can be
CC performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
XX
SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

alignment_scores:
Quality: 2758.00 Length: 525
Ratio: 5.263 Gaps: 1
Percent Similarity: 99.810 Percent Identity: 99.619

alignment_block:
US-09-171-553B-4 x AAT74883
Align seg 1/1 to: AAT74883 from: 1 to: 7393

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpH 17
|||||
598 ATGGGAGGAGGAGCTTAAATTCGAAATTAATCTGCTGTTAAGGCANT 647
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
648 TGAAGTAGATCCAGGCGTCATATTTGTCACTGACGTGAAGAGGAGC 697
34 rTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
698 CTTCGACAGACTTCTGCTCGATGCAATGCGCAACATTCATGTGATG 747
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla1 67
|||||
748 CCAATCAGAGGAGGAGCTTAAATTCGAAATTAATCTGCTGTTAAGGCANT 797
67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84
|||||
798 CATTTTCAAGACTGAGACCGGCGTCATCATTGATCGAGGCCCTAATCC 847
84 eutThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
|||||
848 TTACGTGGCAGAAATTGGCAGAGATCCCGCATGAGGCTTAACCATG 897
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyVal 117
|||||
898 CTAAATTAACCAAGAAAGCCAGGTCCCGAATCTGCTCTTGAGAGAA 947
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuProA 134
|||||
948 AAACAAACACTCGCGCGGAAAAAGTCAGGCCGC .TCCTGCTATCAACCC 996
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
|||||
997 GAGATCGAGAGAGCGCGGACTTGGCGGAACCCCAACCTGTTCCGCCACC 1046
151 ProLeuSerSerThrGlyCysGlyGlyThr.SerAlaProProGlyA 167
|||||

```

1047 CCTTATCCAGCACAGGCTGTGTGAGGGACCTCTGCCCTCCTGGAG 1096  
 167 LAPROVALValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
 1097 CTCGGGTGGTGGAGGAGCCTGCTGCCGGGACTCGGAGCCGGAGAGGCCCC 1146  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG1 200  
 1147 ACCCCGAGGAGACAGACAGATCGGATATTACCGCTGCCACCTATGG 1196  
 200 yProPomMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheS 217  
 1197 CCTCCCATGCGAGGGGGCCAAATTCAGCCCTCCAGATATGGCCCTTTT 1246  
 217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233  
 1247 CTTTCGCAATCTCTATATTGAAAACTAACCATCCCTTTCTCGGAG 1296  
 234 AspProGluArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG1 250  
 1297 GATCCCAAGCCCTCAGCGGTTGGTGGAGTCCCTTATGTTCTCTCACCA 1346  
 250 nProThrTrpAspAspGlyGlnGlnLeuGlnThrLeuPheThrTrg 267  
 1347 GCCTACTTGGAGATGTCAACACGCTCTGCGACACTCTTCACACACCG 1396  
 267 LUG1ArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
 1397 AGGAGCGAGAGAGATTCGTAGAGGCTAAAAAAATGTTCTCTGGGCC 1446  
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProle 300  
 1447 GACGGGCGACCCAGCAGTTCGCAAAATGAGATTGACATGGGATTTCCCTT 1496  
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317  
 1497 GACTGCGCCCGGTGGACTACACACGCGTGAAGGAGGAGACCTTGA 1546  
 317 yS11eTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 1547 AAATCTATTCGCCAGGCTCTGGTGGGGGCTCTCCGGGGCGCTCAAGAGGG 1596  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 1597 CCCACTAATTGGCTAAGTAAGAGAGTGATGACGAGCAGCAAGCAACCC 1646  
 350 OProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP 367  
 1647 TCCTCGGTATTCTTGGAGAGGCTCATGGAAGCCCTTCAGGGGTTCCACC 1696  
 367 rOPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 1697 CTTTGGATCTTAACCTCAGAGGCCCAAGAAAGCTCAGTGGCCCTGCTTC 1746  
 384 l1eGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluG1 400  
 1747 ATTTGGGAGTGGGCTCTGATATCAGGAAGAACTTCAGACACTCGGAAG 1796  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValT 417  
 1797 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAGAGAGCAAGAGAGGTGT 1846  
 417 YTYTARGARGluuThrGluGluGluGluGluGluGluGluGluGlu 433  
 1847 ATTACAGAGAGGAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1896  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 1897 AG 1946  
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467  
 1947 TAAAGCTTGGCCGCGAGTGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 1996

467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 1997 GAGATTTTAGGAAATTAAGTACAGCCCTAGACAGTCAAGGAGACCTGGCC 2046  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlyLysG1 500  
 2047 AATAGAGACCCACTCGACACAGACAGTGTGCTATTGTAAAGAAAAAG 2096  
 500 yHisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysValL 517  
 2097 ACACGTGGGCAAGAGAACTGCCCAAGAAAGCAAGAGAGAGAGAGAGTCC 2146  
 517 euAlaLeuGluGluAspLysAsp 524  
 2147 TAGCTTAGAAGAAATTAAGAT 2169  
 seq\_name: /sids2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67021  
 seq\_documentation\_block:  
 ID AAC67021 standard; DNA; 6076 BP.  
 AC AAC67021;  
 XX  
 DT 27-MAR-2001 (first entry)  
 XX  
 DE PERV env protein coding sequence SEQ ID NO: 21.  
 XX  
 KW Xenotransplantation; infectious agent; vaccine; ds.  
 XX  
 OS Porcine endogenous retrovirus.  
 XX  
 PN WO200071726-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 24-MAY-2000; 2000MO-US14296.  
 XX  
 PR 24-MAY-1999; 99US-0135631.  
 XX  
 PA (MAYO-) MAYO MEDICAL VENTURES.  
 XX  
 PI Federspiel MJ;  
 XX  
 DR WPI; 2001-032041/04.  
 XX  
 PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 PS Claim 16; Page 107-109; 144pp; English.  
 XX  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;  
 alignment\_scores:  
 Quality: 2659.00 Length: 525  
 Ratio: 5.143 Gaps: 1  
 Percent Similarity: 98.476 Percent Identity: 95.619  
 alignment\_block:  
 US-09-171-553b-4 x AAC67021 ..  
 Align seg 1/1 to: AAC67021 from: 1 to: 6076  
 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17

```

1183 ATGGGACAGACGGTGACACCCCTTAGTTGACTCTCCAGACCTTGGAC 1232
17  TGLuValArgSerArgAlaHisAsnLeuSerValGluValLysLysGly 34
1233 TGAAGTTAAATCCAGGGCTCAATTTCTGCTCAGTTAAGGTAAGAGGAC 1282
34  roTPgInThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
1283 CTTGGCCACACTTCTGTGTCTGTAATGGCCGACATTGATGTGGATGG 1332
51  ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
1333 CCATCAGAGGGGACCTTAAATTCGAGATATCCTGGCTGTTAAAGCAAT 1382
67  eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84
1383 TATTTTTCAGACTGGACCCGGCTCTCATCCCAATCAGAGCCCTATATCC 1432
84  euThrTrpGlnAspLeuAlaGluAspProProProTyrValLysProTyr 100
1433 TTACGTGGCAAGATTGGCAGAGATCTCCGCCATGGGTAAACCTTGG 1482
101  LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
1483 CTGAATTAAGCCCAAGAAAGCCAGTCCCGCAATTCGGCTTGGAGAGAA 1532
117  sAsnLysHisSerAlaGlyLysValGluProSerSerSerTyrLeuPro 134
1533 AAACAAACACTCGCTGAAAAAGTCAGCCCTC.TCCTCATATCTATACCC 1581
134  rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
1582 GAGATTGAGAGACCCCGGCTTGGCGGGAACCCCAATCTGTCCGCCACC 1631
151  ProLeuSerSerThrGlyCysGlyLysLysThr.SerAlaProGlyAla 167
1632 CCTTATCTGGCACAGGCTGCGCAGAGGGGACCTCTGCCCTCTCGAG 1681
167  lAProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
1682 CTCGCGGGGTGGAGGACCTGCTGCAAGGACTCGAGCGGGAGGGGCC 1731
184  ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200
1732 ACCCGGAGCGGACAGACGAGATTCGCAATTAACCGTGGCGAGTACG 1781
200  yProPheMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
1782 CCCCCTCCACACCGGGGGCCAAATTCACCCCTCCAGATTTGGCCCTTT 1831
217  eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
1832 CTTTGCACATCTCTATAATTGAAAACTAACCAATCCCCCTTTCGGAG 1881
234  AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250
1882 GATCCCAACAGCCCTCAGCGGGGTGGTGGAGTCCCTTATGTCTCTCACA 1931
250  nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrTrpG 267
1932 GCCTACTGGGATGATTCACACACCTCTCGACACCTTCACAAACCG 1981
267  lUGlArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283
1982 AGGAGCGAGAGAAATCTGTAGAGCTAGAAAAAATGTTCTCTGGGCC 2031
284  AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300
2032 GACGGCGGACCCAGCGAGTTGCAAAATGAGATTGACATGGGATTTCCCT 2081
300  uThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGluSerLeu 317
|||||
2082 GACTCGCCCGGTTGGACTACAAACAGCGCTGAAGGTAGGAGAGCTTGA 2131
317  ystIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArg 333
2132 AAATCTATCGCCAGGCTGTGTGGCGGCTCCGGGGCCCTCCAAAGACG 2181
334  ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlu 350
2182 CCCACTAATTTGGCTAAGGTAAAGAGATGATGCAGGACCCGATGAAC 2231
350  oProSerValPheLeuGluArgLeuMetGluAlaPheArgPheThrP 367
2232 CCCCTCTGTTCCTTGGAGGCTCTTGGAGGCTTCAGCGGTACACCC 2281
367  roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
2282 CTTTGTATCCCACTCAGAGGCCCAAAAGCCCACTGATGGCTTGGGCTT 2331
384  lIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
2332 ATAGGACAGTCAAGCTTGGATATTAGAAAGAGCTTCAGAGACTGGAGG 2381
400  yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValT 417
2382 GTTACAGAGAGCTGAGTTACGTATCTAGTGAAGAGGACGAGAAAGTAT 2431
417  yTyrArgArgGlyIuThrGluGluGlyGlnGluArgLysGlyLysGlu 433
2432 ATTACAAAGAGGACGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2481
434  ArgGluGluArgGluGluArgAspArgArgArgGlnGlyLysAsnLeu 450
2482 AGAGAGAAAGGAGGAGAAAGCTAATTAACGCAAGAGAAAGAAATTGAC 2531
450  rLysIleLeuAlaAlaValAlaGlyLysSerSerArgGluArgGlu 467
2532 TAAATTTTGGCTGACGAGTGTGAAGGAAACCAATACGAAAGAGACA 2581
467  rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483
2582 GAGATTTTAGGAAATTAGTGAGGCCCTTAGACAGTACAGGAAACCTGGGC 2631
484  AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluSerG 500
2632 AATAGACCCCACTCGACAAAGACCAATGTGCTATTGTAAAGAAAGAG 2681
500  yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517
2682 ACACGTGGCAAGACACTGCCCAAGAAAGGAAACAAAGAGACTGAGTGT 2731
517  euAlaLeuGluGluAspLysAsp 524
2732 TAGCTCTGAAAGAAATTAAGAC 2754

seq_name: /STD2/gcgdata/geneseq/geneseqn/NA2001.DAT:AA67032
seq_documentation_block:
ID   AAC67032 standard; DNA; 6076 BP.
XX
AC   AAC67032;
XX
DT   27-MAR-2001 (first entry)
XX
DE   PERV env protein coding sequence SEQ ID NO: 32.
XX
KM   Xenotransplantation; infectious agent; vaccine; ds.
XX
OS   Porcine endogenous retrovirus.
XX
PN   WO200071726-A1.
XX
PD   30-NOV-2000.
XX

```

```

PF 24-MAY-2000; 2000MO-US14296.
XX
XX 24-MAY-1999; 99US-0135631.
XX
XX (MAYO-) MAYO MEDICAL VENTURES.
XX
XX Federspiel MJ;
XX
XX WPL; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
XX transplant recipients, by introducing recombinant DNA comprising DNA
XX encoding extracellular proteins of the agent into donor cells, such as
XX swine cells -
XX
XX Claim 16; Page 117-119; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
XX infectious agents during xenotransplantation. This involves introducing
XX to donor swine cells a recombinant DNA encoding a peptide fragment from
XX the infectious agent, and then introducing these cells into the
XX transplant recipient.
XX
XX Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;

alignment_scores:
Quality: 2659.00 Length: 525
Ratio: 5.143 Gaps: 1
Percent Similarity: 98.476 Percent Identity: 95.619

alignment_block:
US-09-171-553b-4 x AAC67032 ..

Align seg 1/1 to: AAC67032 from: 1 to: 6076

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrp 17
|||||
1183 ATGGGACAGACGGGTGACGACCCCTCTAGTTGACTCGACCATTTGGAC 1232
17 rgluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
|||||
1233 TGAAGTTAAATCCAGGGCTCATTAATTTGCAAGTTCAGGTTAAGAGGAC 1282
34 rdtGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50
|||||
1283 CTGGGACGCTTCTGTCTCTGTAAGTGGCCGACATTCGATGTTGGATGG 1332
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67
|||||
1333 CCATCAGAGGGGACCTTAAATCTCAGATTAATCCCTGGCTGTTAAAGCAAT 1382
67 ellePheGlnThrGlyProGlySerHisProAspGlnLysProTyrIle 84
|||||
1383 TATTTTTCAGACTGGACCGGGCTCTCATCCCAATCAGAGGCCATATATCC 1432
84 eutThrTrpGlnAspLeuAlaGluAspProProProTyrValLysProTrp 100
|||||
1433 TTACGTGGCAGATTTGGCAGAGATCTCCGCCATGGGTTAAACCTTGG 1482
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGlu 117
|||||
1483 CTGATTAAGCCAGAAAGCCAGGTCCCGCATTTCCGGCTCTGGAGAGAA 1532
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134
|||||
1533 AAACAAACACTCGCTGAAAAAGTCACACCCCTC; TCCCTCATATCTACCC 1581
134 rGAspArgGlyAlaAlaAspLeuAlaGluIleThrProThrCysSerProThr 150
|||||
1582 GAGATTTAGAGACCCCGGCTTGGCCGGAACCCCAATCTGTTCGCCACCC 1631
151 ProLeuSerSerThrGlyCysGluGluGlyThrSerAlaProProGly 167
|||||
1632 CCCTATCTGACAGGAGGTGCTGGAGGGAGACCTCTCCCTCTCTGGAG 1681
167 laProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183
|||||
1682 CTCCGGCGGTGGAGGAGCTCTGCGAGGACTCGGAGCGGAGGAGGCGCC 1731
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuAlaGlyTrpGly 200
|||||
1732 ACCCGGAGCGAGACAGAGATCGCAGATTACCGCTGCGCACGTACGG 1781
200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217
|||||
1782 CCCTCCACACCGGGGGGCCAATTGACGCCCTCCAGATTGGCCCTTTT 1831
217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerGlu 233
|||||
1832 CTTCGACAGATCTCTATATTGAAAACTAACCAATCCCTTCTCGGAG 1881
234 AsProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisGly 250
|||||
1882 GATCCCAACCGCTTCACGGGGTGTGGAGTCCCTTATGTTCTCTCACA 1931
250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThr 267
|||||
1932 GCCTACTTGGCATGATTTGTCAACAGCTGCTGCAGACACTCTTCAACA 1981
267 lGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283
|||||
1982 AGGAGCGAGAGAGATTCGTGTAGAGGCTACAAAAATGTTCCCGGGGCC 2031
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
|||||
2032 GACGGGGACCCACGAGTTGCAAAATGAGATTGACATGAGATTTCCTT 2081
300 uThrArgProGlyTrpAspTrpAsnThrAlaGluArgGlnSerLeu 317
|||||
2082 GACTCGCCCGGCTTGGAGTCAACACGCGCTGAAAGTGAAGGAGGCTTGA 2131
317 ysIleTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333
|||||
2132 AAATCTATTCGACAGCTCTGTGGCGGCTCTCCGGGGCGCTCAAGACGG 2181
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350
|||||
2182 CCCACTAATTTGGCTAAGTAAGAGAAAGTATGACGAGGACCGAATGAAC 2231
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367
|||||
2232 CCCTCTGTGTTTCTTGAAGAGGCTTGGAAAGCCTTACGGCGGTACACC 2281
367 rPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
|||||
2282 CTTTGTATCCACCTCAGAGGCCCAAAAGCCATCAGTGGCTTTGGCTTT 2331
384 lIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
|||||
2332 ATAGACAGTACAGGCTTGATATTAGAAAGAGCTTACAGACTGGAAG 2381
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417
|||||
2382 GTTACAGAGAGGCTGAGTTACGTGATCTAGTGAAGAGGCGAGAAAGTAT 2431
417 yTyrArgArgGluThrGluGluLysGlnGlnArgLysGlnLysGlu 433
|||||
2432 ATTACAAAGAGGAGAGAGAAAGAAAGGACAAAGAAAGAAAGAGAGAGA 2481
434 ArgGluGluArgGluGluArgAspArgArgGlnGlnLysAsnLeuTh 450
|||||
2482 AGAGAGAAAGGAGGAAAGACGTAATAACGCAAGAGAAAGTAATTGAC 2531
450 rLysIleLeuAlaValValGluGlyLysSerSerArgGluArgGlu 467
|||||

```

2532 TAAGATCTTGCTGCTGAGTGTGAAGGGAAGAAACAAATACGGAAAGAGA 2581  
 467 rGAspPharGlyLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
 2582 GAGATTTTGAAGAAATTAAGTACAGCCCTAGACAGTCAAGGAGACCTGGGC 2631  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGln 500  
 2632 AATAGACCCCACTGACAGAGACCAATGTGCATATTTGAAGAAAAAG 2681  
 500 yHISrPalAarGAsnCysProLysGlyAsnLysGlyProLysVal 517  
 2682 ACACGGGGCAAGGAGACTGCCCAAGAAAGAAAGAAAGAAAGAGTCT 2731  
 517 euAlaLeuGluGluAspLysAsp 524  
 2732 TAGCTCTGGAAGAGATTAAGAC 2754  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67023

## seq\_documentation\_block:

ID AAC67023 standard; DNA: 7873 BP.  
 AC AAC67023;  
 DT 27-MAR-2001 (first entry)  
 DE PERV env protein coding sequence SEQ ID NO: 23.  
 KW Xenotransplantation; infectious agent; vaccine; ds.  
 OS Porcine endogenous retrovirus.  
 PN W0200071726-A1.  
 PD 30-NOV-2000.  
 PE 24-MAY-2000; 2000WO-US14296.  
 PF 24-MAY-1999; 99US-0135631.  
 PR (MAYO-) MAYO MEDICAL VENTURES.  
 PA  
 PI Federspiel MJ;  
 DR WPI; 2001-032041/04.  
 XX Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 PS Claim 16; Page 112-115; 144pp; English.  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 SO Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

## alignment\_scores:

Quality: 2659.00 Length: 525  
 Ratio: 5.143 Gaps: 1  
 Percent Similarity: 98.476 Percent Identity: 95.619

## alignment\_block:

US-09-171-553b-4 x AAC67023 ..  
 Align seg 1/1 to: AAC67023 from: 1 to: 7873

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpTh 17  
 62 ATGGAGACAGACGGGTGACAGACCCCTTAACTTGAAGTCTGACCATTTGAG 111  
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34  
 112 TGAAGTAAATCCAGGCTCATATATTGTGAGTTCAGGTTAAGAAAGGAC 161  
 34 rOTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
 162 CTGGGAGAGATTCTGTGTCTGAATGGCCGACATTCGATGTGATGATG 211  
 51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAla 67  
 212 CCATCAGAGGGGACCTTAAATCTGAGATTATCTGCTGCTGCTTAAACAA 261  
 67 eLlIephGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
 262 TATTTTCAGACTGGACCGGCTCTCATCCCAATCAGAGACCTTATATCC 311  
 84 eUThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100  
 312 TTACGTGCAAGATTGGCAGAGAGATCTCCGCAATGGGTTAAACCTTGG 361  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117  
 362 CTGAATTAAGCCCAAGAAAGCCAGGTCCCGAATTCCTGCTTGGAGAGAA 411  
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134  
 412 AAACAAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATATCTACCC 460  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 461 GAGATTGAGAGACCGCGGCTTGCGCGAACCCTTCTTCCGCCACCC 510  
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGly 167  
 511 CCTTATCTGCGACAGGCTGCTGCGAGGGAGCCCTGCTCCCTCTGAG 560  
 167 lAproValAlaGluGlyProAlaAlaGlyThrArgSerArgArgAla 183  
 561 CTCGGGCGGTGAGAGACTGCTGCGAGAGACTGGAGCCGAGGGCGCC 610  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyr 200  
 611 ACCCGGAGCGAGACAGAGATCGGACATTAACCTGCGGACGACG 660  
 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpPhe 217  
 661 CCCCTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATTTGGCCCTTT 710  
 217 eRSerAlaAspLeuTyrAsnTrpLysThrAsnHisProPheSerGlu 233  
 711 CTTTGCAGATCTCTATAATTGGAACCTAACCATCCCTTTCTCGGAG 760  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHis 250  
 761 GATCCCAACAGCCCTCAGCGGGTTGTGAGTCCCTATGTTCTCTCACCA 810  
 250 nProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThr 267  
 811 GCCCTACTGGGATGATGTCACAGCTGCTGACAGACCTCTGCACACCG 860  
 267 lGluArgGluArgIleLeuLeuGlnLaaArgLysAsnValProGlyAla 283  
 861 AGGAGCGAGAGAGATTTCTGTAGAGGCTGAAAAAATGTTCCTGGGGCC 910  
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetClyPhePro 300  
 911 GACGGCGAGCCACGACGATTGCAAAATGAGATTGACATGAGATTTCCTT 960  
 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeu 317

```

1010 961 GACTCCCGCCGGTGGACTACACAGCGCTGAAGCTAGCGGAGACTTGA
333 317 ysllelyrArglnAlaValAlaGlyLeuArgGlyAlaSerArgArg
1060 1011 AATCTATGCGCCAGGCTCTGTGGCGGGCTCTCCGGGGCCCTCAAGAGCG
350 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr
1110 1061 CCCACTAATTTGGTAAGSTAAGAGAGATGATGACGAGGCGCAATGAAGCC
367 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThrP
1111 1111 CCCCTGCTTTCTTGAAGAGGCTCTTGGAAACCTTCAGCGCGTACACC
367 367 ropheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe
1161 1161 CTTTGTATGCCACCTCAGAGGCGCCAAAAGCTCAGTGCGCTTTGGCTTTT
384 1161 lIeGlyGlnSerAlaLeuAspPheArgLysLysLeuGlnArgLeuGlu
1211 1211 ATAGGACAGTCAGCCTTGATATTAGAAAGAACTTCAGAGACTGCAAGG
400 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValIT
1261 1261 GTTACAGAGGCTGAGTTACGTGATCTAGTAAGAGGCGCAGAGAAAGTAT
417 417 yTYrArgArgGluThrGlnGluGluLysGlnArgLysGluLysGlu
1311 1311 ATTACAAAAGGAGACAGACAGAGAAAGGAAACAAAGAAAAGAGAGAA
434 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh
1361 1361 AGAGAGGAAAGGAGGAAAGAGCTATATTAACGCGCAAGAGAAATTTCAC
450 450 rLysLleLeuAlaAlaValAlaGluGlyLysSerSerArgLysArgGlu
1411 1411 TTAGATCTTGGCTGCAGTGTGAGGAAAGACATACGGAAGAGAGAGA
467 467 rGAspPheArgLysLleArgSerGlyProArgGlnSerGlyAsnLeuGly
1461 1461 GAGATTTAGGAAATTAGTCAGGCGCTAGACAGTCAGGGAACCTGGCC
484 484 AsnArgThrProLeuAspLysAspGlnCysAlaIleCysLysGluLysG
1511 1511 AATAGAGCCCACTCGACAGAGCCCAATGTGCATATTGTAAGAAAAAGG
500 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal
1561 1561 ACACGGGCAAGGAACTGCCCAAGAGGAAACAAAGAGACTGAAGGTCT
517 517 euAlaLeuGluGluAspLysAsp 524
1611 1611 TAGCTCTGGAAGAGATTAAGAC 1633
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/MA2001.DAT.AAC67019
seq_documentation_block:
ID AAC67019 standard: DNA; 7362 BP.
XX
XX AAC67019;
XX
XX 27-MAR-2001 (first entry)
XX
XX PERV env protein coding sequence SEQ ID NO: 19.
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
XX Porcine endogenous retrovirus.
XX
XX WO200071726-A1.
XX
XX 30-NOV-2000.
PD

```

```

XX
XX 24-MAY-2000; 2000MO-US14296.
XX
XX 24-MAY-1999; 99US-0135631.
XX
XX (MAYO-) MAYO MEDICAL VENTURES.
XX
XX Federspiel MJ;
XX
XX MPI; 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
XX transplant recipients, by introducing recombinant DNA comprising DNA
XX encoding extracellular proteins of the agent into donor cells, such as
XX swine cells -
XX
XX Claim 16; Page 101-104; 144pp; English.
XX
XX The present invention provides a method to prevent the transmission of
XX infectious agents during xenotransplantation. This involves introducing
XX to donor swine cells a recombinant DNA encoding a peptide fragment from
XX the infectious agent, and then introducing these cells into the
XX transplant recipient.
XX
XX Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;
XX

```

```

alignment_scores:
Quality: 2655.00 Length: 525
Ratio: 5.135 Gaps: 1
Percent Similarity: 98.476 Percent Identity: 95.429

```

```

alignment_block:
us-09-171-553b-4 x AAC67019

```

```

Align seg 1/1 to: AAC67019 from: 1 to: 7362

```

```

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpH 17
1248 ATGGAGACAGCGGTGACGACCCCTCTAGTTTACTCTGCACCATTTGGAC 1297
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
1298 TGAAGTTAAATCCAGGCGCTCATATTTGTACGTTCAAGATTAAAGAGGAC 1347
34 roTrpGlnThrPheCysAlaSerGluThrProThrPheAspValGlyTrp 50
1348 CTGGCCAGACTTCTGTCTGTGAATGGCCGACATTCGATGTTGGATGG 1397
51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
1398 CCATCAGAGGGGACCTTAAATCTGAGATTATCTGCGTGTAAACCAAT 1447
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrllel 84
1448 TATTTTTCAGACCTGAGCCCGCTCATCCCAATCAGAGCCCTATATATCC 1497
84 euthTrpGlnAspLeuAlaGluAspProProThrProValLysProTrp 100
1498 TTACGTGGCAAGATTGGCCAGAGATCTCCGCCATGGGTTAAACCTTGG 1547
101 LeuAsnLysProArgLysProGlyProArgLysIleLeuAlaLeuGlyL 117
1548 CTGAATTAAGCCAGAAAGCCAGGTCCCGCAATTTCTGCTTGGAGAGAA 1597
117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrlleuPro 134
1598 AAACAAACACTCGCTGGAAGAAAGTCAAGCCCTC.TCCTCATATCTAACCCC 1646
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150
1647 GAGATTGAGAGAGCGCGGCTTGGCCGGAACCCCAATCTGTGTTCCGCCACC 1696

```

	seq_name:	/SID52/gcgdata/geneseq/geneseq/NAI97.DAT:AAT74884
seq_documentation_block:	ID	AAT74884 standard; cDNA; 7892 bp.
XX AC AAT74884;		
XX DT 09-FEB-1998	(first entry)	
XX DE Miniatute swine retrovirus CDNA.		
KM Retrovirus; porcine; GAG protein; POL protein; ENV protein; KM xenotransplantation; infectious; provirus; organ transplant; donor activated virus; PCR; ss. XX OS Porcine retrovirus.		
XX Key Location/Qualifiers		
FH CDS 585..2159		
FT /*tag= a		
FT mat_peptide 585..2156		
FT /*tag= b		
FT /note= "putative GAG protein"		
FT CDS 2307..5744		
FT /*tag= c		
FT mat_peptide 2307..5741		
FT /*tag= d		
FT /note= "putative POL protein"		
CDS 5620..7536		
FT /*tag= e		
FT mat_peptide 5620..7533		
FT /*tag= f		
FT /note= "putative ENV protein"		
XN WC9721836-AI.		
XN PN 19-JUN-1997.		
XP PD 13-DEC-1996;	96WO-US19680.	
XP PR 14-DEC-1995;	95US-O572645.	
PA (GEHO ) GEN HOSPITAL CORP.		
PI Fishman JA;		
DRE WPI; 1997-332804/30.		
DR P-PADB; AAW32096-W32098.		
P New nucleic acid from porcine retroviruses - used for detecting PT viruses in transplant or other tissue and for assessing risk of PT transmitting infection to graft recipient		
PS Claim 22; Fig 3; 128bp; English.		



XX This cDNA sequence represents a porcine retrovirus from miniature swine  
 CC containing the putative coding regions for viral GAG, POL and ENV  
 CC proteins. This sequence and PCR fragments generated from it  
 CC (see AAT74812-T74882) can be used to screen organs for the presence of  
 CC porcine retroviruses prior to xenotransplantation. Transplantation can  
 CC increase the likelihood of retroviral activation if intact and  
 CC infectious proviruses are present. The porcine retroviral sequence can be  
 CC used to generate probes to determine the level (e.g. copy number) of  
 CC intact (i.e. potentially replicating) porcine provirus sequences in a  
 CC strain of xenograft transplantation donors. It can be used to detect  
 CC mutations, genetic lesions or viral recombinants and also to determine  
 CC the histological localisation of activated retroviruses. Using Polymerase  
 CC Chain Reaction DNA quantitation (PQD) on blood mononuclear cells,  
 CC infectivity titration and susceptibility testing can be performed.  
 CC Ultimately animal donors without intact porcine retroviral sequences or a  
 CC lower copy number of viral elements could be selected.  
 SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

## alignment\_scores:

Quality: 2640.00 Length: 525  
 Ratio: 5.116 Gaps: 1  
 Percent Similarity: 98.286 Percent Identity: 94.476

## alignment\_block:

US-09-171-553b-4 x AAT74884 ..

Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrTrp 17  
 |||||||  
 585 ATGGGACAGACGGTGCACACCCCTTGTAGTTGACCTGACCATTTGGAC 634  
 17 rgluValArgSerArgAlaHisLeuSerValGlnValLysLysGly 34  
 |||||||  
 635 TGAAGTTAAATCCAGGGGCTCAATATTGTCAGTTCAGTTAAGAGGAG 684  
 34 rGTrpGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
 |||||||  
 685 CTGGCAGACTTTCGTGTCTGATGATGGCCGACATTCGATGTTGGATGG 734  
 51 ProSerGluGlyThrPheAsnSerLulleleuAlaValLysAlaIle 67  
 |||||||  
 735 CCATGAGAGGGGACCTTAATTCAGATATTCCTGGCTTAAAGCAGT 784  
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProtyrIle 84  
 |||||||  
 785 TATTTTCAGACTGACCCGGCTCTCATCCGATCAGAGGCCATATATCC 834  
 84 eutThrTrpGlnAspLeuAlaGluAspProProProThrValLysProTr 100  
 |||||||  
 835 TTACGTGCAGAGATTGGCAGAGATCTCCGCCATGGGTTAAACCAAG 884  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117  
 |||||||  
 885 CTGAATTAAGCCAGAAAGCCAGAGTCCCGAATTCGTGGCTTGGAGAGAA 934  
 117 sasnLysHisSerAlaGluLysValGluProSerSerSerLysLeuPro 134  
 |||||||  
 935 AAACAAACACTCGCTGAAAAAGTCACGCCCTC.TCCATCATATCTACCCC 983  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 |||||||  
 984 GAGATTGAGAGACCAACCGCTTGGCCGGAACCCCAATCTGTCCCCACAC 1033  
 151 ProLeuSerSerThrGlyCysCysGlyLysThr.SerAlaProProGly 167  
 |||||||  
 1034 CCTTATCTGACAGGGGTGGCGGAGGAGCCCTTTGGCCCTCTCTGAG 1083  
 167 laproValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
 |||||||

1084 CTCGGCGGTGGAGGACCTGCTGCAGGAGACTCGAGGCCGGAAGGCGCC 1133  
 184 ThrProGluArgThrAspGlnIleAlaIleLeuProLeuArgThrTrp 200  
 |||||||  
 1134 ACCCGGAGCGGACAGACAGATCGGACATTACCGCTGGCAGCTACGAG 1183  
 200 yProPheMetProGlyLysGlnLeuGlnProLeuGlnIleThrTrpPhe 217  
 |||||||  
 1184 CCTTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATGGCCCTTTT 1233  
 217 eSerAlaAspLeuThrAsnTrpLysThrAsnHisProPheSerGlu 233  
 |||||||  
 1234 CTTCGAGATCTCTATATTGAAAACTAACATCCCTCTTCTCGGAG 1283  
 234 AspProGlnArgLeuThrGlyLeuValGlnLysLeuMetPheSerHisG 250  
 |||||||  
 1284 GATCCCAACGCCCTCCACGGGGTGGTGAGCCCTTATGTTCTCTCACA 1333  
 250 nProThrTrpAspCysGlnGlnLeuLeuGlnThrLeuPheThrTrp 267  
 |||||||  
 1334 CCTTACTTGGGATGATTTCTCAACAGCTGCTCAGACACTTTCACAA 1383  
 267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
 |||||||  
 1384 AGGACCGAGAGAGATTTCTATTAGAGGCTACAAAAATGTTCCCTGG 1433  
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300  
 |||||||  
 1434 GACGGGCAACCCACCGCGTTCGCAAAATGAGATTCATGAGATTTCCCT 1483  
 300 uThrArgProGlyTrpAspTrpAsnThrAlaGlnGlyArgGlnSerLeu 317  
 |||||||  
 1484 AACTCCGCCCGGTGGGACTCAACACGCTCGAAGGTAAGGAGAGCTTGA 1533  
 317 ySileThrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 |||||||  
 1534 AAATCTATCGCCAGAGCTCTGTTGGCGGCTCCGGGCGCCTCAAGACG 1583  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnLysProAsnGlu 350  
 |||||||  
 1584 CCCCTAATTTGGCTAAGGTAAAGTAAGTATGAGGAGCCGGAATGACC 1633  
 350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgArgPheThr 367  
 |||||||  
 1634 CCCCTGTTTCTTCTGAGAGGCTCTTGAAGCCTTCAGCGGATACACC 1683  
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 |||||||  
 1684 CTTTGTATCCCACTTCAGAGGCCCAAAAGCCCTCAGTGGCTTGGCCTT 1733  
 384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400  
 |||||||  
 1734 ATAGGACAGTCAAGCTTGGATATTAGAAAGAACCTTCAGAGCTGGAAG 1783  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417  
 |||||||  
 1784 GTTACAGAGAGGTGAGTTAGTACGTATGATGAGGAGGCGAGAAAGTAT 1833  
 417 yTrpArgArgGluThrGlnGluGlnLysGlnLysArgLysGluLysGlu 433  
 |||||||  
 1834 ATTAACAAAAGGAGACGAAAGAAAGGAAAGGAAAGAAAGGAGAGAA 1883  
 434 ArgGluGluArgGluGluArgAspArgArgGlnGluLysAsnLeu 450  
 |||||||  
 1884 AGAGAGAAAGGAGAGAAAGACCTAATAAAGGCAAGAGAGAAATTGAC 1933  
 450 rLysIleLeuAlaAlaValValGluLysSerSerArgArgGluAla 467  
 |||||||  
 1934 TAAGATCTTGCTGCAAGTGGTTGAAGGAAACCAATAGGAAAGAGAGA 1983  
 467 rGAspPheArgLysIleArgSerGlyProArgGlnSerLysAsnLeuGly 483  
 |||||||  
 1984 GAGATTTTAAAGAAATTAGTCAAGCCCTAAGACAGTCAAGGAAACCTGG 2033

484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGln 500  
 2034 AATGAGACCCACTCGACAGACCAATGTCATATTGTAAAGAAAGAG 2083  
 500 YHISTPRLAARGAsnCysProLysLysGlnLysGlnLysValL 517  
 2084 ACACCTGGGCAAGAACTGCCCCGAAAGGAAACAAAGACCAAGATCC 2133  
 517 euAlaLeuGluGluAspLysAsp 524  
 2134 TAGCTCTAGAGAAAGATTAAGAT 2156  
 seq\_name: /SID52/gcseqdata/geneseq/geneseqn/NA2001.DAT:AAF77727

seq\_documentation\_block:  
 ID AAF77727 standard; DNA; 8132 BP.

XX AAF77727;  
 XX 23-MAY-2001 (first entry)  
 XX Nucleotide sequence of a retrovirus found in miniature swine.  
 XX Retrovirus; graft transplantation; xenotransplantation; miniature swine;  
 XX SS.  
 XX Unidentified.  
 XX US6190861-B1.  
 XX 20-FEB-2001.  
 XX 13-DEC-1996; 96US-0766528.  
 XX 14-DEC-1995; 95US-0572645.  
 XX (GEHO) GEN HOSPITAL CORP.  
 XX Fishman JA;  
 XX WPI; 2001-256211/26.  
 XX P-PSDB; AAB73285, AAB73286, AAB73287.  
 XX Assessing risk of endogenous retroviruses in clinical practice and in  
 PT xenotransplantation, comprises using probe sequences derived from swine  
 PT or miniature swine retroviral genome -  
 XX Claim 1; Fig 3; 127pp; English.  
 XX The present invention relates to a method for screening a cell or tissue  
 CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g. the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.  
 XX Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment\_scores:  
 Quality: 2640.00 Length: 525  
 Ratio: 5.116 Gaps: 1  
 Percent Similarity: 98.286 Percent Identity: 94.476

alignment\_block:  
 US-09-171-553b-4 x AAF77727 ..

Align seq 1/1 to: AAF77727 from: 1 to: 8132

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpTh 17  
 585 ATGGGACAGACCGTGAACGACCCCTTATGTTGACTCTCGACCATTTGAC 634  
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysLysGlyP 34  
 635 TGAAGTTAATCCAGGCGCATTAATTGTGAGTTGAGTTGAAGAGGAC 684  
 34 rotRPolInThrPhcCysAlaSerGluTrrProThrPheAspValGlyTrp 50  
 685 CTGGGACAGACTTCTGTCTCTGATGCGCACATTCGATGATGATGATG 734  
 51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAlaI 67  
 735 CCATCGAGGGAGCCTTATTTCTGACATTAATCTGCGCTTTAAAGCACT 784  
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleL 84  
 785 TATTTTTCAGACTGGACCGGCTCATCCGATCCGATCAGAGCGCTTATCC 834  
 84 euThrTrpGlnAspLeuAlaGluAspProProProTrrValLysProThr 100  
 835 TTACGTGGCAAGATTTGGCAGAGATCTCCCGCATGGGTTAAACCATGG 884  
 101 LeuAsnLysProArgLysProGlyProArgLLeuAlaLeuGlyLysL 117  
 885 CTGCAATTAACCAAGAAAGCCAGGTCCCGCATTCGGCTTGGAGAGAA 934  
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTrrLeuProA 134  
 935 AAACAAACACTCGGCTGAAAAAGTCAAGCCCTC.TCCTCATTCATACCC 983  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 984 GAGATTGAGAGACCGCGGCTTGCGCGGAGAACCCCATCTGTTCCCCAC 1033  
 151 ProLeuSerSerThrGlyCysCysGluGlyThr.SerAlaProProGlyA 167  
 1034 CCTTATTCGGGACAGGTGCGCGGAGGACCTTGGCCCTCGGAG 1083  
 167 laProValValGluGlyProAlaAlaGlyThrArgSerArgGlyAla 183  
 1084 CTCGGCGGGGAGGACCTGTCAGGAGCTCGGAGCGGAGGAGGCGCC 1133  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
 1134 ACCCCGAGCGGACAGACAGATCGGACATTACCGCTGCGACAGTACGG 1183  
 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTrrTrpProPhe 217  
 1184 CCCCTCCACACCGGGGGCCCAATTCAGAGCCCTCCAGTATGCGCTTTT 1233  
 217 eSerAlaAspLeuTrrAsnTrpLysThrAsnHisProProPheSerGlu 233  
 1234 CTCTGAGATCTCTATTAATGGAAAACTACCATCCCTTCTTGAGG 1283  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250  
 1284 GATCCCAACAGCCCTCAGCGGGGTGGAGTCCCTTATGTTCTCACCACA 1333  
 250 nProThrTrpAspAspCysGlnGlnLeuGlnInThrLeuPheThrThg 267  
 1334 GCCTACTTGGGATGATGATTCACACAGCTGTCAGACACTCTTACACACG 1383  
 267 LuGluArgGluArgIleLeuLeuGlnAlaArgLysAsnValProGlyAla 283  
 1384 AGGAGCCAGAGAGATTCATTTAGAGCTTGAAAAAATGTTCTCTGGGGCC 1433  
 284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300  
 1434 GACGGGCGACCCACAGCGGCTTGCAAAATGAGATTGACATGGGATTTCCCTT 1483

```

300 uThArpProGlyTTrpAspTyrAsnThrAlaGluGlyAArgGluSerLeuL 317
1484 AACTGCGCCCGTTGGAGCTACACACGCGCTGAAGAGTAGAGAGCTTGA 1533
317 ySlleTArpGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArpArg 333
1534 AAATCTATCGCCAGCTCTGTGGGGTCTCCGGGGCGCTCAAGACGG 1583
334 ProThrAsnLeuAlaLysValAlaGluValMetClnGlyProAsnGluPr 350
1584 CCCACTATTTGGCTTAAGTAAGAGAGATGATGACGAGGACCGAATGAC 1633
350 oProSerValPheLeuGluArgLeuMetGluAlaPheArpArgPheThrP 367
1634 CCCCTCTGTTTCTTCTGAGAGGCTCTTGAAGCCTTCAGCGGTAACCC 1683
367 ropheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383
1684 CTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGGCTTGGCTT 1733
384 lLeGlyInSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
1734 ATAGGACAGTCAGCCTTGATATTAGAAAGACCTTCAGAGACTGGAGG 1783
400 yLeuGlnGluAlaGluLeuArgAspLeuValAlaArgGluAlaGluValT 417
1784 GTTACAGAGAGCTGAGTTACGTGATCTAGTGAAGAGAGCGACAAAGTAT 1833
417 yTTrpArpArgGluThrGluGluGluGluGluGluGluGluGluGlu 433
1834 ATTACAAAGGGAGACAGAAAGAAAGAAAGAAAGAAAGAAAGAGACAA 1883
434 ArgGluGluArgGluGluArgAspArgGluGluGluGluGluGluGlu 450
1884 AGAGAGAAAGGGAGAGAGAGCTATATAAGCGCAAGAGAGAGATTGGAC 1933
450 rlySlleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467
1934 TAAGATCTTGGCTGCAAGTGTGTAAGGAAAGCAATACGCAAGAGAGA 1983
467 rGAspPheArpGlyLysIleArgSerGlyProArgIleSerLysLeuGly 483
1984 GAATTTTGGAAATTTAGGTCAAGGCCCTAGACAGTCAAGGAAACCTGGGC 2033
484 AsnArpThProLeuAspLysAspGlnCysAlaTyrCysLysGluGlu 500
2034 AATAGGACCCCACTCGACAAAGACCAATGTGCATATTGTAAGAAAGAG 2083
500 yHlSTpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValL 517
2084 ACACTGGGCAAGAACTGCCCAAGAGGAAAGAAAGAGCAAGAGATCC 2133
517 euAlaLeuGlnGluAspLysasp 524
2134 TACCTTAGAAGAAAGATTAAAGAT 2156

```

seq\_name: /SIDS2/gcdata/geneseq/geneseq/NA1997.DAT: AAT74811

seq\_documentation\_block:

```

ID AAT74811 standard; cDNA; 8060 BP.
XX
AC AAT74811;
XX
DT 11-FEB-1998 (first entry)
XX
DE Porcine retrovirus Tsukuba-1 cDNA.
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; Tsukuba-1; PCR; ss.
XX
OS Porcine retrovirus.

```

```

XX Key Location/Qualifiers
FH 86..2002
FT CDS /*tag= a
FT /*note= "putative ENV protein"
FT 3112..4686
FT /*tag= b
FT /*note= "putative GAG protein"
FT 4871..8060
FT CDS /*tag= c
FT /*note= "putative POL protein (partial)"
XX
XX WO921836-A1.
XX
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX
XX 14-DEC-1995; 95US-0572645.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Fishman JA:
XX
XX WPI: 1997-332804/30.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX
XX Claim 1; Fig 1; 128pp; English.
XX

```

This sequence represents the purified porcine retroviral cDNA sequence of Tsukuba-1 and contains the putative coding regions for viral proteins GAG, POL and ENV. This sequence and PCR fragments generated from the sequence (see AAT74812-T74882) could be used to screen organs for porcine retroviruses prior to xenotransplantation. Transplantation can increase the likelihood of retroviral activation if intact and infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantation donors. It can be used to detect mutations, genetic lesions or viral recombinants and to determine the histological localization of activated retroviruses. Using polymerase chain reaction DNA quantitation (PQO) on blood mononuclear cells, infectivity titration and susceptibility testing can be performed. Ultimately animal donors without intact porcine retroviral sequences or with a lower copy number of viral elements could be selected.

Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment\_scores:

```

Quality: 2634.00 Length: 525
Ratio: 5.105 Gaps: 1
Percent Similarity: 98.286 Percent Identity: 94.286

```

alignment\_block:

US-09-171-553B-4 x AAT74811 ..

Align seg 1/1 to: AAT74811 from: 1 to: 8060

```

1 MetGlyGlnPheValThrThrProLeuSerLeuThrLeuAspHisTyrPth 17
3112 ATGGGACAGACGGTGCAGACACCCCTTGAAGTTGACTCTCGACCAATGGAC 3161
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
3162 TGAAGTTAAATCCAGGCGCTCATATTTGTCAGTTCAAGTTAAGAGGAGC 3211
34 rortPpGlnThrPheCysAlaSerGluThrProThrPheAspValGlyTyr 50
3212 CTGGCAGACTTCTGTCTGTGAATGGCCGACATTGATGTTGGATGG 3261

```

51 ProSerGluGlyThrPheAsnSerGluIleLeuAlaValLysAlaIle 67  
 |||||  
 3262 CCATCAAGAGGGGACCTTAAATTCAGATATATCTGCTGTAAAGCAAT 3311  
 |||||  
 67 eilePheGluThrGlyProGlySerHisProAspGluGluProTyrIle 84  
 |||||  
 3312 TATTTTCAGACTGACCGGGCTCTCATCCGATCAGAGAGCCCTATATCC 3361  
 |||||  
 84 eutHTrpGlnAspLeuAlaGluAspProProTyrValLysProTyr 100  
 |||||  
 3362 TTACGTGGCAAGATTGGCAGAGATCTCCGCATGGTTAAACATG 3411  
 |||||  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117  
 |||||  
 3412 CTGAATAGCCAGAAAGCCAGGTCCCGAATTCGTGCTTGAGAGAA 3461  
 |||||  
 117 SASLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134  
 |||||  
 3462 AAACAAACACTCGGCTGAAAAAGTCMAAGCCCTC.TCCTCATATCTACCC 3510  
 |||||  
 134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
 |||||  
 3511 GAGATTGAGAAACACCGGCTTGGCCGGAACCCCAATCTGTCCGCCACC 3560  
 |||||  
 151 ProLeuSerSerThrGlyCysCysGluGluThr.SerAlaProProGly 167  
 |||||  
 3561 CCTTATCTGGACAGGGGTGCCGAGAGGACCTTGTGCCCTCTCTGAG 3610  
 |||||  
 167 lApProValGluGluProAlaAlaGlyThrArgSerArgArgGlyAla 183  
 |||||  
 3611 CTCGGCGGTGGAGGACCTTCTGCAGGACTCGGAGCCGAGGAGGCC 3660  
 |||||  
 184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
 |||||  
 3661 ACCCGGAGCGGACAGAGATCGGCACATTACCGCTGCCACAGTACGG 3710  
 |||||  
 200 yProMetProGlyGlyGluLeuGluProLeuGlnTyrTyrProPhe 217  
 |||||  
 3711 CCTCTCCACACCGGGGGGCCAATTGCACCCCTCCAGTATTTGCCCTTT 3760  
 |||||  
 217 eSerAlaAspLeuTyrAsnTyrLysThrAsnHisProProPheSerG 233  
 |||||  
 3761 CTTTGCAGATCTATATATTTGAAACTAAACATCCCTTCTTCGGAG 3810  
 |||||  
 234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250  
 |||||  
 3811 GATCCCAACGCTCAGCGGGGTGGAGTCCCTTATGTTCTCTCACCA 3860  
 |||||  
 250 nProThrTrpAspAspCysGlnGluLeuLeuGlnThrLeuPheThr 267  
 |||||  
 3861 GCCTACTTGAGATTTGTCAACACGCTGTCAGACACTCTTCACACCG 3910  
 |||||  
 267 lGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
 |||||  
 3911 AGGAGCGAAGAGAAATTCATTTAGAGGCTAGAAAAAATGTTCTCTGGGCC 3960  
 |||||  
 284 AspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPhePro 300  
 |||||  
 3961 GACGGGCGACCCACGCGGTTGCAAAATGATTTGACATGGATTTCCTT 4010  
 |||||  
 300 uThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGluSerLeu 317  
 |||||  
 4011 AACTGCCCGCGGTGGACTACACACGCTGAGAGTAGGAGGAGACTTGA 4060  
 |||||  
 317 ySLeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArg 333  
 |||||  
 4061 AAATCATGCGCAGGCTCTGTGGCGGCTCTCCGGGCGCCCTCAAGACGG 4110  
 |||||  
 334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGluPr 350  
 |||||  
 4111 CCCACTAATTTGGCTAAGGTAAAGAGATGATGACGAGGCCAATGAAACC 4160

350 oProSerValPheLeuGluArgLeuMetGluAlaPheArgPheThrP 367  
 |||||  
 4161 CCCCTTGTTTTTCTTGAGAGGCTCTTGAACCTTCAAGCGGTACACC 4210  
 |||||  
 367 roPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
 |||||  
 4211 CTTTGTATCCCACTCAGAGGCCCAAAAGCCTCATGCTTGTGGCTTT 4260  
 |||||  
 384 lIeGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuG 400  
 |||||  
 4261 ATAGGACAGTCAGCCTTGATATTAGAAAGAAAGCTTACAGACTGGAAG 4310  
 |||||  
 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGlyValT 417  
 |||||  
 4311 GTTACAGAGAGCGTGAATTACGTGATCTAGTAAGAGGACAGAAATAT 4360  
 |||||  
 417 yTTrArgArgGluThrGluGluGluGluGluArgGlySerGlyGlu 433  
 |||||  
 4361 ATTACAAAAGCGAGACGAGAAAGAAAGGAAACAAAGAAAGAGAGAA 4410  
 |||||  
 434 ArgGluGluArgGluGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
 |||||  
 4411 AGAGAGAAAGGAGAGAAAGCCTATTAACGGCAAGAGAAATTTGAC 4460  
 |||||  
 450 rLysIleLeuAlaAlaValAlaGluGlyLysSerSerArgGluArgGlu 467  
 |||||  
 4461 TAAGATCTTGCTGTCAGTGTGTAAGGAAAGCAATACGAAAGAGAA 4510  
 |||||  
 467 rGAspPheArgGlyLysSerGlyProArgGlnSerGlyAsnLeuGly 483  
 |||||  
 4511 GAGATTTTGAAGAAATTAAGTCAAGGCCCTTAGACAGTCAGGAAACCTGGGC 4560  
 |||||  
 484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG 500  
 |||||  
 4561 AATAGAGACCCACTCGACAGAGACCAATGTCATATTGTAAGAAAGAGG 4610  
 |||||  
 500 yHisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysVal 517  
 |||||  
 4611 ACACTGGGCAAGAAACTGCCCAAGAGGAAAGCAAGAACCAAGGATCC 4660  
 |||||  
 517 euAlaLeuGluGluAspLysAsp 524  
 |||||  
 4661 TAGCTCTAGAAGAAAGATAAAGAT 4683  
 |||||  
 seq\_name: /SID2/gc9data/geneseq/geneseqn/NA2001.DAT:AAF77725  
 seq\_documentation\_block:  
 ID AAF77725 strand:..cDNA: 8060 BP.  
 XX AAF77725;  
 AC  
 XX  
 DT 23-MAY-2001 (first entry)  
 DE Tsukuba-1 cDNA.  
 XX  
 KW Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.  
 XX  
 OS Porcine retrovirus.  
 XX  
 PN US6190861-B1.  
 XX  
 PD 20-FEB-2001.  
 XX  
 PF 13-DEC-1996; 96US-0766528.  
 XX  
 PR 14-DEC-1995; 95US-0572645.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Fishman JA;  
 XX  
 DR WPI; 2001-256211/26.  
 XX

PT Assessing risk of endogenous retroviruses in clinical practice and in  
PT xenotransplantation, comprises using probe sequences derived from swine  
PT or miniature swine retroviral genome -  
PS  
XX Claim 1; Fig 1; 127pp; English.  
XX  
CC The present invention relates to a method for screening a cell or tissue  
CC for the presence or expression of a retrovirus (RV), comprising  
CC contacting a target nucleic acid from the cell or tissue with a second  
CC nucleic acid from the present invention (e.g. the present sequence or a  
CC fragment thereof). The method is useful for RV detection and to assess  
CC graft transplantation risk. Screening of animals allows the elimination  
CC of donors with active replication of known viruses. Inactive proviruses  
CC can be detected and inactivated, allowing identification and elimination  
CC of potential human pathogens derived from swine in a manner not possible  
CC in the outbred human organ donor population and is important to the  
CC development of human xenotransplantation.

Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment\_scores:  
Quality: 2634.00 Length: 525  
Ratio: 5.105 Gaps: 1  
Percent Similarity: 98.286 Percent Identity: 94.286

alignment\_block:  
US-09-171-553B-4 x AAF77725 ..

Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrPth 17  
|||||  
3112 ATGGGACAGAGCGTGTGAGACCCCTTGTGAGTCTCGACCATTTGGAC 3161  
17 TGIuValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34  
|||||  
3162 TGAAGTTAAATCCAGGCGCTCAATTTGTGAGTTCAGGTTAAGAGGAC 3211  
34 roTrGlnThrPheCysAlaSerGluTrpProThrPheAspValGlyTrp 50  
|||||  
3212 CTTGGCAGACTTCTGTCTGTGATGCGGACATTCGATGATGGATGG 3261  
51 ProSerGlnGlyThrPheAsnSerGluLeuLeuAlaValLysAlaI 67  
|||||  
3262 CCATAGAGGGGACCTTAATTCAGATTAATCCGCGCTGTTAAGCAT 3311  
67 eIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
|||||  
3312 TATTTTTCAGACTGAGCCGCTCTCATCCGATCAGAGGCCCTATATCC 3361  
84 euThrTrGlnAspLeuAlaGluAspProProPthProValLysProTrp 100  
|||||  
3362 TTACGTGGCAGGATTTGGCAGAGATCTCCGCCATGGGTTAAACCATGG 3411  
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGly 117  
|||||  
3412 CTGATTAAGCCAGAAAGCCAGGTCGCCGAAATTCGCTTGGAGAA 3461  
117 sAsnLysHisSerAlaGlyLysValGluProSerSerSerTyrLeuPro 134  
|||||  
3462 AAACAAACACTCGCTGAAAAAGTCAGCCCTC; TCCATATATCTACCCC 3510  
134 rGAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSerProThr 150  
|||||  
3511 GAGATTGAGGAACACCGCGCTTGGCGGAGACCCCATCTGTCCGCCACC 3560  
151 ProLeuSerSerThrGlyCysCysGlnGlyThr; SerAlaProProGly 167  
|||||  
3561 CCTTATCTGCGACAGGGTGGCGGAGGCCCTTGTGCCCTCTCTGAG 3610  
167 laProValValGlnGlyProAlaAlaGlyThrArgSerArgArgGlyAla 183  
|||||

3611 CTCGGCGGTGAGGAGACTTTCAGAGGACTCGGAGCCGGAGGCCGCC 3660  
184 ThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrG 200  
|||||  
3661 ACCCGGAGCGGACAGAGATCGGACATTCACCGCTGGCGACGTACGG 3710  
200 yProPheMetProGlyGlnLeuGlnProLeuGlnTyrTrpProPhe 217  
|||||  
3711 CCTCCACACCGGGGGCCCAATTCAGAGCCCTCCAGTATTCGCTTTT 3760  
217 eSerAlaAspLeuTyrAsnTrpLysThrAsnHisProProPheSerG 233  
|||||  
3761 CTTCTGAGATCTCTAATATGGAACAACTAACCATCCCTTCTCTGGAG 3810  
234 AspProGlnArgLeuThrGlyLeuValGluSerLeuMetPheSerHisG 250  
|||||  
3811 GATCCCAACGCGCTCAGCGGCTTGTGTGAGTCCCTATATGTTCTCTC 3860  
250 nProThrTrpAspAspCysGlnGlnLeuGlnThrLeuPheThrThr 267  
|||||  
3861 GCTACTTGGGATGATGTCACACAGCTGTCAGACACTCTCTCAACACG 3910  
267 lueGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAla 283  
|||||  
3911 AGGAGCGAGAGAAATTCATTAAGAGCTAAGAAAAATGTTCTGGGGCC 3960  
284 AspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPhePro 300  
|||||  
3961 GACGGCGGACCCACCGCGTTCGCAAAATGAGATTGACATGGGATTTCC 4010  
300 uThrArgProGlyTyrPaspTyrAsnThrAlaGluGlyArgGluSerLeu 317  
|||||  
4011 AACTCGCCCGGTGGGACTACAAACGCGCTGAAGTAGAGGAGAGCTTGA 4060  
317 ySIIeTyrArgGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArg 333  
|||||  
4061 AAATCTATCCGACGCTGCTGCGGCTCCGGGCGCTCAACACGG 4110  
334 ProThrAsnLeuAlaLysValArgGluValMetGlnGlyProAsnGlu 350  
|||||  
4111 CCCACTATTTGGCTAAGTAAGAGAAATGATCAGAGCCGATGATGACC 4160  
350 oProSerValPheLeuGlnLysLeuMetGluAlaPheArgArgPheThr 367  
|||||  
4161 CCCCTGTTTCTTCTGAGAGGCTCTGGAAGCTTCAGCGGCTACACCC 4210  
367 roPheAspProThrSerGlnAlaGlnLysAlaSerValAlaLeuAlaPhe 383  
|||||  
4211 CTTTTCATCCCACTCAGAGGCCCAAAAGCTCAGTGGCTTGTGGCTTT 4260  
384 IleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400  
|||||  
4261 ATAGAGACAGTACGCTTGATATTAGAAAGAACTTCAGAGACTGGAGG 4310  
400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysVal 417  
|||||  
4311 GTTACAGAGAGCTGAGTTACGTATCTAGTGAAGGAGCGCAGAAAGTAT 4360  
417 yTyrArgArgGluThrGlnGluGluLysGlnArgLysGluLysGlu 433  
|||||  
4361 ATTACAAAGAGGAGACAGAAAGAAAGGAGCAAAAGAAAGAGAGGAA 4410  
434 ArgGlnGluArgGluArgArgAspArgArgGlnGluLysAsnLeuTh 450  
|||||  
4411 AGAGAGAAAGGAGGAGAAAGCGTAATTAACGCGCAAGAAATTTGAC 4460  
440 rLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluArgGlu 467  
|||||  
4461 TAAGATTTGCTGCACTGTTAAGGAGAAACCAATACGAAAGAGAGAG 4510  
467 rGAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGly 483  
|||||  
4511 GAGATTTTAGGAAATTAAGTCAAGGCCCTAGACAGTCAGGAGAACCTG 4560

484 AsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysG1 500  
 4561 AATAGGACCCCACTCGACAAAGACCAATCTCATATTGTAAGAAAGAG 4610  
 500 yHISTPAlaArgAsnCysProLysLysGlnLysGlyProLysValL 517  
 4611 ACACCTGGCAAGAACTGCCCAAGAAAGGAAACAAAGAGCAAGATCC 4660  
 517 euAlaLeuGluGluAspLysAsp 524  
 4661 TAGCTCTAGAGAGATTAAGAT 4683  
 seq\_name: /SID52/9cdata/geneseq/geneseq/NA1999.DAT:AAV69750  
 seq\_documentation\_block:  
 ID AAV69750 standard; cDNA; 8655 BP.  
 XX AAV69750;  
 XX  
 DT 01-MAR-1999 (first entry)  
 XX  
 DE Mus dunni endogenous virus.  
 XX  
 KW MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;  
 KW vector; S8.  
 XX  
 OS Mus dunni endogenous virus.  
 XX  
 XX Key location/Qualifiers  
 FT CDS 422..2290  
 FT /tag= a  
 FT /transl\_except= (pos:1..2, aa:Met)  
 FT /product= Gag (glycosylated)  
 FT CDS 722..2290  
 FT /tag= b  
 FT /product= Gag  
 FT 2291..5902  
 FT /tag= c  
 FT /product= Pol  
 FT CDS 5754..7775  
 FT /tag= d  
 FT /product= Env  
 XX  
 PN W09850538-A1.  
 XX  
 PD 12-NOV-1998.  
 XX  
 PF 08-MAY-1998; 98WO-US09452.  
 XX  
 PR 08-MAY-1998; 98US-0075272.  
 PR 09-MAY-1997; 97US-0046140.  
 XX  
 PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
 XX  
 PI Bonham L, Miller AD, Wolgamot G;  
 XX  
 DR WPI: 1999-034718/03.  
 DR P-PSDB: AAM81570-73.  
 XX  
 PT New retroviral packaging cells - containing Mus dunni endogenous  
 PT virus sequences to target cells, retrovirus gag and pol genes and a  
 PT heterologous gene of interest.  
 XX  
 PS Claim 3; Page 60-64; 85pp; English.  
 XX  
 CC This is the complete nucleotide sequence of Mus dunni endogenous  
 CC virus (MDEV) including open reading frames encoding glycosylated  
 CC gag (see W815700, Gag (see AAM81571), Pol (see AAM81572) and Env (see  
 CC AAM81573). A cultured packaging cell is claimed which produces a  
 CC replication-defective retroviral vector (RDV) particle, where  
 CC the packaging cell is a vertebrate cell capable of expressing and  
 CC assembling retroviral proteins, comprising: (a) a first vector

CC encoding a retroviral envelope protein having amino acid residues  
 CC MDEV that direct binding of the retroviral particle to MDEV  
 CC retroviral receptors on a target cell; and (b) a second vector  
 CC encoding retrovirus gag and pol proteins, where upon expression of  
 CC the vectors in the packaging cell in the presence of a vector  
 CC having a sequence of a heterologous gene of interest, a  
 CC replication-defective retroviral particle is produced that binds to  
 CC MDEV receptors of target cells. Also claimed are: (1) a cultured  
 CC packaging cell for producing a RDV particle; (2) methods for  
 CC producing a RDV particle comprising a heterologous gene of  
 CC interest; (3) cultured packaging cell line PD22; and (4) a RDV  
 CC produced by a method as in (2). The MDEV receptor is present on a  
 CC variety of cells rendering MDEV pseudovirus packaging cells useful  
 CC in methods of mammalian and particularly human gene transfer for  
 CC gene therapy. The MDEV packaging cells are a stable and  
 CC reproducible source of retroviral particles. Clones may be  
 CC isolated from these populations that produce high titre virus. The  
 CC packaging cell lines may be selected and cloned for other desirable  
 CC properties, such as stability of in vivo growth, lack of production  
 CC of helper virus, lack of reinfection by viral particles packaged in  
 CC the cell, stability from genetic rearrangement and recombinational  
 CC events, resistance to complement lysis, and improved ability to  
 CC infect cells from higher mammals.

XX S0 Sequence 8655 BP; 2316 A; 2228 C; 2051 G; 2060 T; 0 other;

alignment\_scores:  
 Quality: 1684.00 Length: 541  
 Ratio: 3.907 Gaps: 10  
 Percent Similarity: 79.667 Percent Identity: 61.553

alignment\_block:  
 US-09-171-553B-4 x AAV69750 ..

Align seg 1/1 to: AAV69750 from: 1 to: 8655

1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisThrPth 17  
 722 ATGGGACAACTCTCTCCACGCTCCCTTCTCTAACCCTGGAGCATGGAA 771  
 17 GlnValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34  
 772 GGAGTAAATAATCAGACACACACATCATGCTGAGTGAAGAGGCC 821  
 34 roTrpGlnThrPheCysAlaSerGlnTrpProThrPheAspValGlyTrp 50  
 822 CATGGCAAACTTTTGGCGCTCCGAGTGGCAACGTTTGGAGTGGCTGG 871  
 51 ProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67  
 872 CCACCGGAGGGTGTCTTTGACTTGTCTCATGATCGCCGCCCTCAGGCAAT 921  
 67 IlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIle 84  
 922 TGTTTTCAGGAA...GAAGGGGTCACTCATGATCATCTCCATCATTTG 968  
 84 euThrTrpGlnAspLeuAlaGluAspProProProProValLysProTrp 100  
 969 TGACCTGGCAGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1018  
 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117  
 1019 .....ACCCCAATCTCTGCAAACTGACGCTGCACTGTC 1053  
 117 sAsnLysHisSerAlaGluLysValGluProSerSer.....SerTyrL 132  
 1054 CCAATCCGATGAGCTGAGTGAATCTGTGCTGCTGCTGCTGCTGCTGCT 1103  
 132 euProArgAspArgGlyAlaAlaAspLeuAlaGlyThrProThrCysSer 148  
 1104 ATCCAGAGTATGACGACCTCTCTGTGATGATGATCCCAACCTCCCTTAC 1153



CC retroviruses with cell membranes, while maintaining mutant envelope  
 CC protein incorporation into a virion, and viral titers of about two  
 CC orders of magnitude within that observed for wild-type retrovirus when  
 CC the protein or fragment is expressed on the surface of a retroviral  
 CC particle. The proteins have an increased ability to penetrate targets,  
 CC typically cells and a correspondingly increased ability to deliver  
 CC nucleic acids or drugs. The mutated nucleic acid is useful for gene and  
 CC drug therapy, especially as drug delivery vehicles. The retrovirus  
 CC particles are useful to transduce eukaryotic cells. The retrovirus  
 CC cells are useful in the treatment of cancer in a human. Other diseases  
 CC contemplated for treatment include adenosine deaminase deficiency (ADA),  
 CC thalassemia, hemophilia, diabetes, alpha-anti trypsin deficiency, brain  
 CC and neural disorders, phenylketonuria, growth disorders, heart diseases  
 CC and immune diseases.

XX Sequence 8088 BP; 2103 A; 2251 C; 1982 G; 1752 T; 0 other;

# alignment\_scores:

Quality: 1660.50 Length: 542  
 Ratio: 3.944 Gaps: 10  
 Percent Similarity: 77.675 Percent Identity: 61.439

## alignment\_block:

US-09-171-553B-4 x AA245540 ..

Align seg 1/1 to: AA245540 from: 1 to: 8088

1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisThrPh 17  
 631 ATGGGACAAAGATTAATCTACCCCTATCTCCCTCACTTAATCACTGAG 680  
 17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34  
 681 AGATGGACAAACAGGCTCACAATCTATCCCTGGAATCAAAAGGAA 730  
 34 rGTPGlnThrPheCysAlaSerGluTTPProThrPheAspValGlyT 50  
 731 AATGGCAGACTTCTGTTCTCCGAGTGGCCACATCGGCGTGGGGTGG 780  
 51 ProSerGluGlyThrPheAsnSerGluLeuLeuAlaValLysAlaI 67  
 781 CCACGGAGGAGACTTTAATCTCTGTCATTTTGGCATTAAGAT 830  
 67 eilePheGlnThrGlyProGlySerHisProAspGlnGluProTyrLe 84  
 831 TGTCTTCAGGAGAAC..GGGGACATCCGGACCAAGTCCATATACG 877  
 84 eutThrTrpGlnAspLeuAlaGluAspProProTProTValLysProT 100  
 878 TGGTATGGCAGGACCTCGCCCAAAATCCCCACCATGGGTG..... 918  
 101 LeuAsnLysProArgLysProGlyProArgGlyLeuAlaLeuGlyL 117  
 919 .....CCAGCTTCGCCCAAGGTCGCTGTGTCTGATAC 953  
 117 sAsnLysHisSerAlaGluLysValGluProSerSerSerTyrLeuPro 134  
 954 CCGAGACACAGTTCGGGGAGG.....CCATCAGCTCTCCCGACCC 997  
 134 rAspArgGlyAlaAsp.....LeuAlaGlyThrProThr... 146  
 998 CCATCTACCCGACACAGACTTACTCTCTGTGAAACCCACAGGCC 1047  
 147 .....CysSerProThrProLeuSerSerThrGlyCysCys 158  
 1048 CCGCCCTATCCGGCGCACTGCCACCCCTTGCGCCCTCAGGCGATGG 1097  
 158 sGluGlyThrSerAlaProGlyAlaProValAlaGluGlyProAla 175  
 1098 ACCGCGCTCAGGCGACGATGCGGATAGTACGATCCTGAGGGCGACG 1147  
 175 lAgGlyThrArgSerArgArgGlyAlaThrPro.....GluArgThr 188

1148 CTGGAGCAGGAGTCCGCTGCCCGCAGTCCACAGCAACTCGGCTCT 1197  
 189 AspGluLeuAlaIleLeuProLeuArgThrTyrGlyProProMetPro 205  
 1198 GACTCCACTGTGATTTTGGCTCCCGAGCCATAGAGCCCCCGGACGC 1247  
 205 yGlyGlnLeuGlnProLeuGlnIntTyrTrpProPheSerSerAlaSp 222  
 1248 CAATGGCTGTGCTCTTACAAATATGCGCTTTTCTCCAGAGATCTT 1297  
 222 yAsnTrpLysThrAsnHisProProPheSerGluAspProGlnAla 238  
 1298 ATAATGTGAATCTTAATCATCCCTTTTCTGAAAAACAGAGAGTCT 1347  
 239 ThrGlyLeuValGluSerLeuMetPheSerHisGlnProThrTrpAs 255  
 1348 ACGGGCTCTTGAGTCTTATGTTCTCCATCAGCCACTTGGAGCA 1397  
 255 pCysGlnGlnLeuLeuGlnThrLeuPheThrThrGluArgGluArg 272  
 1398 TTGCCAAGAGTCTCTACAGATTTCTTCCACCACTGAGAGCAAGAA 1447  
 272 lLeuLeuGlnAlaArgLysAsnValProGlyAlaAspGlyArgPro 288  
 1448 TTCCTCGAGGCGCCGCAAAATGCTTGGGACAAATGGGGCCCTACA 1497  
 289 GlnLeuGlnAsnGluLeuAspMetGlyPheProLeuThrArgProG 305  
 1498 CAGCTCCAGAACCTTAAATGAGGCTTCCCTCAATGAGACTACTG 1547  
 305 pAspTyrAsnThrAlaGluArgLysLeuLysIleTyrArgGln 322  
 1548 GGATTACAAACACAGCCGAGGTAGAGGAGCTCTTCTGTACCGCGG 1597  
 322 lAlaValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeu 338  
 1598 CTCTAGGCGAGCTCAAAAGGGGAGCTGGGCTCTTACCAATTTGG 1647  
 339 LysValArgGluValMetGlnGlyProAsnGluProProSerValP 355  
 1648 AAGGTAAAGAGAGTCTTCCAGGGACCGGACAGAACCCCTCGTT 1697  
 355 uGluArgLeuMetGluAlaPheArgArgPheThrProPheAspPro 372  
 1698 AGAAGCGCTGATGGAGGCTTATAGAGATACACTCCCTTTGATCC 1747  
 372 eArgGluAlaGlnLysAlaSerValAlaLeuAlaPheIleGlyIns 388  
 1748 CTGAGGAGACACAGGCTGGCTGCGCATGCTTTATCGGACAGTCA 1797  
 389 LeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnAla 405  
 1798 CCAGATATCAAGAAAGATTACAGAGCTAGAGGGCTCCAGAGCTAT 1847  
 405 uLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 422  
 1848 CTTCAAGATTATTAAGAGGCGCAAAAGGTGTACCAATTAAGACAG 1897  
 422 hArgGluGluLysGluGlnArgLysGluLysGluArgGluArgGlu 438  
 1898 CAGAAAGAGAGACAGAAAGAGAAAGAAAGAGGACAGAAAGAGAG 1947  
 439 GluArgArgAspArgGlnGluLysAsnLeuThrLysIleLeuAla 455  
 1948 AGCGGGCGGCGATAGCCGAGAAAGAAACCTTAAATTCCTGGCGG 1997  
 455 aValValGluGlyLysSerSerArgGluArgGluArgAspPheArg 472  
 1998 AGTAGTA.....AGTAGA..... 2013  
 472 lArgSerGlyProArgGlnSerGlyAsnLeuGlyAsn..... 484



```

2014 ..GGGTCCACAGTAGGACAGACAGGAACTGAGCACACAGGCAAGAG 2061
485 .....ArgThrProLeuAspLysAspGlnGysAlaIleTyrC 496
2062 AACCTAGGATGAGAAAGCTCTCACTAGACAAAGACAGTCGCGCTACTG 2111
496 sLysGlnLysGlnIleThrPalaIleArgAsnCysProLysLysGlnAsn...L 512
2112 TAAAGAGAAAGGGCCATGTGGCAAGAGATGTCCCGAATAAACACGCTCA 2161
512 ysgLysProLysValLeuAlaLeuGlu 520
2162 GAGAGCCAGGATCTTACGCTAGAT 2187
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ73731
seq_documentation_block:
ID AAQ73731 standard; DNA: 8535 BP.
XX
AC AAQ73731:
XX
DT 10-MAY-1995 (first entry)
XX
DE GALV SEATO genome.
XX
KM GALV; gibbon-ape-leukemia virus; retro virus; vector; plasmid p558;
KW plasmid p521; plasmid p537; transgenic animal; gene therapy; ss.
XX
OS Gibbon leukemia virus SEATO.
XX
PN WO9423048-A.
XX
PD 13-OCT-1994.
XX
PF 06-APR-1994: 94MO-US03784.
XX
PR 06-APR-1993: 93US-0043311.
XX
PA (USSH ) US SEC DEPT HEALTH.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Deacon NJ, Eiden MV, Hooker DJ, Wilson CA.
XX
DR WPT: 1994-333206/41.
XX
PT Recombinant DNA contg. defective gibbon ape leukemia virus
PT genome - and specific gene, also derived mammalian cells, virions
PT etc., useful in gene therapy of e.g. cancer, viral infections or
PT inherited diseases
XX
PS Disclosure; Page 30-34; 68pp; English.
XX
CC Novel replication-deficient retro virus vectors p558 (AAQ73733), p521
CC (AAQ73734) and p537 (AAQ73735) include a GalV-derived packaging site,
CC esp. the 200-910 base region of the GalV SEATO genome, and
CC regulatory sequences, partic. the GalV 3' long terminal repeat
CC (AAQ73732). The vectors allow transfer of genes for gene therapy and
CC transgenic animal breeding.
XX
SQ Sequence 8535 BP; 2210 A; 2376 C; 2084 G; 1865 T; 0 other:

```

## alignment\_scores:

```

Quality: 1622.50      Length: 538
Ratio: 3.836          Gaps: 11
Percent Similarity: 78.625      Percent Identity: 60.595

```

## alignment\_block:

```
US-09-171-553B-4 x AAQ73731 ..
```

```
Align seg 1/1 to: AAQ73731 from: 1 to: 8535
```

```
1 MetGlnLthrValThrThrProLeuSerLeuThrLeuAspHisThrPh 17
```

```

||||| ..... |||||
1001 ATGGGACACAGATATTCTACCCCTATCTCCCTCACTTATATCATCTGGAG 1050
17 rGluValArgSerArgAlaHisAsnLeuSerValGlnValLysGlyP 34
1051 AGATGTGAGAAACAAGGCGCTCACAATCTATCCGTGAAATCAAAAAGGAA 1100
34 rorTpgLthrThrPheCysAlaSerGlnThrProThrPheAspValGlyTyr 50
1101 AATGGCAGACTTCTGTCTCTCCAGTGGCCACATTCGGCGGGGGTGG 1150
51 ProSerGlnLysThrPheAsnSerGluIleIleLeuAlaValLysAlaI 67
1151 CCAOCGGAGGGAAGCTTTTATCTCTCTGTCATTTTGGCACTTAAAAAGAT 1200
67 eIlePheGlnThrGlyProGlySerHisProAspGlnLysProTyrIle 84
1201 TGTCTTTCAGAGAGAAC...GGGGACATCCGGACCAAGTTCATATATCG 1247
84 eurThrTrpGlnAspLeuAlaGluAspProProThrTrpValLysProTyr 100
1248 TGGTATGGCAGAGACCTCGCCAGAAATCCCAACATGGGTG..... 1288
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLys 117
1289 .....CCAGCTCCGGCAAGGTCGCTGTCTGTCTGTATAC 1323
117 sAsnLysHisSerAlaGlnLysValGluProSerSerSerTyrLeuPro 134
1324 CCGAGACCAAGTTCGGCGAGG.....CCATAGCTCTCCCGCCAGCCCC 1367
134 rgaSparGlyAlaAlaAsp.....LeuAlaGlyThrProThr... 146
1368 CCATCTACCCGGACACAGAGACTTACTCTCTCTCTCAACCCACAGCCCC 1417
147 .....CysSerProThrProLeuSerSerThrGlyCysC 158
1418 CCGCCCTATCCGGCGCAGCTCGCCAGCCCTCGCCCTCAGCGCATCGG 1467
158 sGlnGlyThrSerAlaProProGlyAlaProValAlaGlnGlyProAla 175
1468 ACCGCCGTACGGCCAGATGCGCATGTAGCGATCTCGAGGGCCAGCGG 1517
175 laGlyThrArgSerArgArgAlaThrPro.....GluArgThr 188
1518 CTGGACACAGAGATCCCGCGTCCCGCAGTCCAGACAGACACTCGGTCCT 1567
189 AspGlnIleAlaIleLeuProLeuArgThrTyrGlyProProMetProG 205
1568 GACTCCACGTGATTTTGGCCCTCCGAGCCATAGACACCCCGCGAGCC 1617
205 yGlnGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAspLeu 222
1618 CAATGGCTGTGTCCTCTCAATATATGAGCCCTTTTCTCAGAGATCTTT 1667
222 yTrsAsnTrpLysThrAsnHisProProPheSerGlnAspProGlnArgLeu 238
1668 ATATATTGGAATCTATCATCTCCCTCTTTTCTGAAAACACAGAGGTCTC 1717
239 ThrGlyLeuValGlnSerLeuMetPheSerHisGlnProThrTrpAsp 255
1718 ACGGCGCTCTTGTAGTCTCTATGTCTCCATCAGACCCACATGGAGCA 1767
255 pCysGlnGlnLeuLeuGlnThrLeuPheThrThrGlnGluArgLysI 272
1768 TTGCCACACAGCTCTTCAAGATTCCTTTCACCATGAGCAACGGGAAGA 1817
272 leuLeuGlnLalaArgLysAsnValProGlyAlaAspGlyArgProThr 288
1818 TTCTCTCGAGGCGCCCAAAATGTCTTGGGAGACAATGGGCGCCCTACA 1867
289 GlnLeuGlnAsnGlnIleAspMetGlyPheProLeuThrArgProGlyTyr 305
||||| ..... |||||

```

1868 CAGCTCGAGACCTCATTAATGAGCCTCCCTCAATCGACCTCACTG 1917  
305 PASPTyAsnThrAlaGluGlyArgGluSerLeuLysIleTyArgGlnA 322  
1918 GGATTACACACACCGAGGTAGGAGCGCTTCTGGTCTACCGCCGGA 1967  
322 IAlaValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAla 338  
1968 CTCTAGTGGCAGGTCTCAAGGGCAGCTCGCGCTCTACCAATTTGGCT 2017  
339 LysValArgGlyValMetGlnGlyProAsnGluProProSerValPheLe 355  
2018 AAGGTAAAGAGGTCTTGTACAGGACCGCAGAACCCCTTCGGTTTCTT 2067  
355 uGluArgLeuMetGluAlaPheArgArgPheThrProPheAspProThrs 372  
2068 AGAAGCGCTGATGGAGCGCTATGAGAGATACACTCCGTTGATCCCTCTT 2117  
372 eRGluAlaGlnLysAlaSerValAlaLeuAlaPheIleGlyGlnSerAla 388  
2118 CTGAGGGGCAACAGCGTGGCGCTCCCATGGCTTATCGGACAGTCAAGC 2167  
389 LeuAspIleArgLysLysLeuGlnArgLeuGlnGlyLeuGlnGluAlaG 405  
2168 CCAGATATCAAGAAAGTTACAGAGGCTAGAGGGGCTCCAGGACTATTC 2217  
405 uLeuArgAspLeuValArgGluAlaGlnLysValTyrTyrArgArgGluT 422  
2218 CTTACAAAGATTAGTAAAGAGCGCAAAAGGTGTACCATATAGAGAGAGA 2267  
422 hrgIuGluGluLysGluGlnArgLysGluLysGluArgGluArgGlu 438  
2268 CAGAAAGAGAAAGACAAAGAAAGAAAAAGAGAGCAAGAAAGAGAG 2317  
439 GluArgArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaI 455  
2318 AGGCGGCGCATAGGCCGAGAGAAAAAACTTGACTAAATTTGGCCGC 2367  
455 aValVal.....GluGlyLysSerSerArgGluArgGluArgAspPheA 470  
2368 AGTAGTAGTAGAGAGAGGCTCCACAGGTAGCGACCCCTCCCGGA.... 2413  
470 rGlySIIeArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnArgThr 486  
2414 .....CCTAGGGTAACTGACTGTG...GAGGGGAGCC 2443  
487 ProLeuAsp.....LysAspGlnCysAlaTyrCysLysGluLysGln 500  
2444 CCCATTGAGTCCCTGTCAAAGACCAAGTCCCATCTGTATAAGAGAGGG 2493  
500 yHISTrpAlaArgAsnCysProLysLysGlyAsn...LysGlyProLysV 516  
2494 CCATTGGCGCAGAGATGTCCCGAAAAAACACGTCAAGAGAACCAAGG 2543  
516 aLeuAlaLeuGlu 520  
2544 TTCTAGCCCTAGAT 2557

OM of: US-09-171-553B-5 to: N\_Geneseq\_1101:\* out\_format : pfs  
 Date: Feb 24, 2002 10:19 AM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
 -O=/cgn2.1/USPTO.spool/US09171553/runat\_22022002.122826\_11203/app\_query.fasta\_1.3312  
 -DB=N\_Geneseq\_1101 -OPMT=fastap -SUFFIX=feb22std.rng  
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000  
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DALIGN=200  
 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=0 -MODE=LOCAL  
 -OUTFMT=pfs -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09171553\_CGNL\_1526 -NCPU=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-171-553B-5  
 Query length: 1194  
 Database: N\_Geneseq\_1101:\*  
 Database sequences: 930621  
 Database length: 428662619  
 Search time (sec): 338.200000

## score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SID52/gcgdata/geneseq/NA1997.DAT:AAV09700		6304.00	8311.45	0.0	8209
/SID52/gcgdata/geneseq/NA1997.DAT:AAV09569		6284.00	8284.99	0.0	8196
/SID52/gcgdata/geneseq/NA1997.DAT:AAV09699		6099.50	8041.13	0.0	7892
/SID52/gcgdata/geneseq/NA2001.DAT:AAV77727		6099.50	8041.13	0.0	8132
/SID52/gcgdata/geneseq/NA2001.DAT:AAV670120		6095.50	8036.55	0.0	4402
/SID52/gcgdata/geneseq/NA2001.DAT:AAV67021		6095.50	8036.55	0.0	7362
/SID52/gcgdata/geneseq/NA2001.DAT:AAV67023		5858.00	7721.41	0.0	7873
/SID52/gcgdata/geneseq/NA2001.DAT:AAV67032		5646.50	7444.11	0.0	6076
/SID52/gcgdata/geneseq/NA1997.DAT:AAV74811		5637.50	7429.29	0.0	8060
/SID52/gcgdata/geneseq/NA2001.DAT:AAV77725		5637.50	7429.29	0.0	8060
/SID52/gcgdata/geneseq/NA2001.DAT:AAV67021		5635.50	7429.54	0.0	6076
/SID52/gcgdata/geneseq/NA1997.DAT:AAV09698		4840.00	6382.62	0.0	3320
/SID52/gcgdata/geneseq/NA2001.DAT:AAV77726		4489.50	5910.46	0.0	7333
/SID52/gcgdata/geneseq/NA1997.DAT:AAV74883		4489.50	5910.38	0.0	7393
/SID52/gcgdata/geneseq/NA1999.DAT:AAV69750		4339.50	5710.18	5.5e-310	8655
/SID52/gcgdata/geneseq/NA2000.DAT:AAZ45540		4319.00	5683.74	1.6e-308	8088
/SID52/gcgdata/geneseq/NA1994.DAT:AAQ73731		4319.00	5683.18	1.6e-308	8535
/SID52/gcgdata/geneseq/NA1996.DAT:AAQ94266		4000.50	5261.94	5.1e-285	8502
/SID52/gcgdata/geneseq/NA1995.DAT:AAQ94266		3988.00	5244.56	4.8e-284	8892
/SID52/gcgdata/geneseq/NA1995.DAT:AAQ81042		3980.00	5234.65	1.7e-283	8323
/SID52/gcgdata/geneseq/NA1997.DAT:AAV80058		3980.00	5234.65	1.7e-283	8323
/SID52/gcgdata/geneseq/NA1999.DAT:AAZ07063		3980.00	5234.65	1.7e-283	8323
/SID52/gcgdata/geneseq/NA1999.DAT:AAV80057		3980.00	5232.40	2.3e-283	8367
/SID52/gcgdata/geneseq/NA1999.DAT:AAZ07067		3980.00	5232.40	2.3e-283	10367
/SID52/gcgdata/geneseq/NA1999.DAT:AAV80058		3979.50	5234.05	1.8e-283	8278
/SID52/gcgdata/geneseq/NA2001.DAT:AAV89136		3979.50	5234.05	1.8e-283	8330
/SID52/gcgdata/geneseq/NA1999.DAT:AAV77178		3967.00	5221.41	9.3e-283	8330
/SID52/gcgdata/geneseq/NA2000.DAT:AAV87841		3967.00	5221.41	9.3e-283	8330
/SID52/gcgdata/geneseq/NA1997.DAT:AAV80693		3967.00	5218.78	5.2e-283	7308
/SID52/gcgdata/geneseq/NA1997.DAT:AAV80693		3967.00	5218.78	5.2e-283	7308
/SID52/gcgdata/geneseq/NA1997.DAT:AAV80698		3967.00	5218.36	1.4e-282	7616
/SID52/gcgdata/geneseq/NA1995.DAT:AAV80452		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1996.DAT:AAV07374		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1997.DAT:AAV05061		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1998.DAT:AAV42715		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1998.DAT:AAV19551		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1998.DAT:AAV15346		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA1999.DAT:AAV28143		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA2000.DAT:AAV40297		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA2000.DAT:AAZ50064		3967.00	5217.43	1.5e-282	8332
/SID52/gcgdata/geneseq/NA2000.DAT:AAZ45539		3967.00	5217.43	1.5e-282	8332

/SID52/gcgdata/geneseq/NA1997.DAT:AAV09555 + 3967.00 5213.57 2.5e-282  
 /SID52/gcgdata/geneseq/NA1997.DAT:AAV09558 + 3967.00 5202.23 1.1e-281  
 /SID52/gcgdata/geneseq/NA1994.DAT:AAQ38700 + 3959.00 5206.85 6.0e-282  
 /SID52/gcgdata/geneseq/NA1994.DAT:AAQ38701 + 3958.00 5203.27 9.5e-282  
 /SID52/gcgdata/geneseq/NA2001.DAT:AAV67022 + 3929.50 5173.21 4.5e-280  
 493

seq\_name: /SID52/gcgdata/geneseq/NA1997.DAT:AAV09700

seq\_documentation\_block:

ID AAV09700 standard; DNA; 8209 BP.

XX AAV09700;

XX 20-MAY-1998 (first entry)

XX Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.

XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;

XX virion core polypeptide; polymerase protein; envelope protein;

XX vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT misc\_feature 1..61

FT /tag= a

FT /note= "R LTR domain"

FT LTR 1..588

FT /tag= b

FT misc\_feature 62..143

FT /tag= c

FT /note= "U5 LTR domain"

FT CDS 588..2162

FT /tag= d

FT /product= GAG protein

FT /note= "viral core polypeptide"

FT CDS 2163..5747

FT /tag= e

FT /product= POL protein

FT /note= "polymerase peptide"

FT CDS 5620..7590

FT /tag= f

FT /product= ENV protein

FT /note= "envelope protein"

FT LTR 7591..8209

FT /tag= g

FT misc\_feature 7638..8106

FT /tag= h

FT /note= "U3 LTR domain"

FT protein\_bind 7643..7686

FT /tag= i

FT /note= "Binding site for transcription factors OCT-1, c-Myb, LyF-1, E47 and ETS-1"

FT protein\_bind 7719..7739

FT /tag= j

FT /note= "Binding site for transcription factors AP-4 and NF1"

FT protein\_bind 7744..7754

FT /tag= k

FT /note= "Binding site for transcription factor AP-1/TR"

FT protein\_bind 7773..7785

FT /tag= l

FT /note= "Binding site for transcription factor ETS-1/GATA"

FT protein\_bind 7793..7834

FT /tag= m

FT /note= "Binding site for transcription factors GATA, ETS-1, c-Myb, AP-1, GATA, E47"

FT protein\_bind 7841..7851

FT /tag= n

FT /note= "Binding site for transcription factor AP-1"

FT protein\_bind 7873..7883

FT /tag= o

FT /note= "Binding site for transcription factor AP-1"

protein\_bind 7885..7914  
 /\*tag= p  
 /note= "Binding site for transcription factors  
 AP-1/GATA, c-Myb and AP-1"

protein\_bind 7928..7953  
 /\*tag= q  
 /note= "Binding site for transcription factors B47  
 and ETS-1"

CAAT\_signal 7964..7968  
 /\*tag= z  
 /note= "Binding site for transcription factor GATA"

protein\_bind 7985..7997  
 /\*tag= s  
 /note= "Binding site for transcription factor GATA"

CAAT\_signal 8003..8007  
 /\*tag= t  
 /note= "Binding site for transcription factor GATA"

protein\_bind 8024..8036  
 /\*tag= u  
 /note= "Binding site for transcription factor GATA"

CAAT\_signal 8042..8046  
 /\*tag= v  
 /note= "Binding site for transcription factor GATA"

protein\_bind 8049..8060  
 /\*tag= w  
 /note= "Binding site for transcription factor GATA"

TATA\_signal 8075..8090  
 /\*tag= x  
 /note= "Binding site for transcription factor GATA"

misc\_feature 1..61  
 /\*tag= y  
 /note= "R LTR domain"

misc\_feature 8127..8135  
 /\*tag= z  
 /note= "PolyA downstream element"

polyA\_signal 8161..8166  
 /\*tag= aa

W09740167-A1.  
 30-OCT-1997. 97WO-GB01087.  
 18-APR-1997; 97WO-GB01087.  
 10-FEB-1997; 97GB-0002658.  
 19-APR-1996; 96GB-0008164.  
 (IMUT-) IMUTRAN LTD.  
 (OONE-) Q-ONE BIOTECH LTD.  
 Galbraith DN, Haworth C, Lees GM, Smith KT;  
 WPI; 1997-535851/49.  
 P-PSDB; AAW39271; AAW39272; AAW39273.  
 Polynucleotide encoding porcine retrovirus expression product -  
 useful to develop products for use in vaccines, diagnosis and  
 xeno-transplantation

Claim 4; Fig 3; 69pp; English.

This DNA sequence encodes the porcine retrovirus (POEV) virion core  
 polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and  
 also includes the Long Terminal Repeat (LTR). These proteins can be used  
 to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 antiviral agents. They can also be used in xeno-transplantation  
 technology and as diagnostic tools.

Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

alignment\_scores:  
 Quality: 6304.00 Length: 1194  
 Ratio: 5.280 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-171-553B-5 x AAV09700 ..  
 Align seg 1/1 to: AAV09700 from: 1 to: 8209

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
 2163 GGGAGACGGGGTTCGAGCCCTCCCGAGCCAGGGTAACCTTGAAGGT 2212  
 17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
 2213 GGAAGGGCAACAGTTGAGTCTCTGGTTGATACCGAGCGGAGCATTCAG 2262  
 34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
 2263 TGCTGCTACAACCATAGGAAACTAAAGAAAAAATCCTGGGTGATG 2312  
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67  
 2313 GGTGCCACAGGGCAACGGCAGTATCCATGGACTACCCGAAGAACCTTGA 2362  
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
 2363 CTGGGAGTGGGACGGTAACCCACTCGTTCTGGTTCATCCCTGAGTGCC 2412  
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
 2413 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAA 2462  
 101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
 2463 TCTTTTGACAAGAGAGACCCAGAGGTCTGTGTAATAACAACCCATCAC 2512  
 117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134  
 2513 TGTGTTGACCCCTCCAATTAGATGATGAATATCGACTATATTCTCCCAAG 2562  
 134 allysProAspGlnAspIleGlnSerTrpLeuGlnPheProGlnAla 150  
 2563 TAAAGCCTGATCAAGATATACAGTCTGTGGTGGAGCAGTTTCCCAAGCC 2612  
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
 2613 TGGCAGAAACCCGAGGATGGGTTGGCAAGCAAGTTCCCCACAGGT 2662  
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
 2663 TATTCAACTGAAGGCCAGTGTCTACACCAATATCAGTCAGACAGTACCCCT 2712  
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
 2713 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAGATTAAATC 2762  
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
 2763 CAACAGGGCATCTAGTCTCTGTCCTCAATCCCTTGGAAATACCTCCCTGCT 2812  
 217 uproValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
 2813 ACCGTTTAGGAAGCCTGGGACCAATGATTATCGACCAAGTACAGGACTTGA 2862  
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
 2863 GAGAGGTCAATAAAGGGTGCAGGACATACACCCACGGTCCCGAACCCCT 2912  
 251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
 2913 TATAACCTCTTGAGCGCCCTCCCGCTGACGGAACTGGTACACAGTATT 2962  
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
 2963 GGACTTAAAGATGCCTTCTCTGCTGAGATTACACCCCACTAGCCAAAC 3012

284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
 |||||  
 3013 CACTTTTGGCTTCGAATGGAGAGATCCAGTACGGGAAGAACCGGCAG 3062  
 |||||  
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
 |||||  
 3063 CTCACCTGGACCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3112  
 |||||  
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
 |||||  
 3113 TGAGGAAGCCCTACACAGGACCTGGCCAACTTCAGAGATCCAAACCCCTC 3162  
 |||||  
 334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
 |||||  
 3163 AGGTGACCTCCCTCCAGTACGTGGATGACCTGCTTCTGGGGGAGCCACC 3212  
 |||||  
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
 |||||  
 3213 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTTGCTGA 3262  
 |||||  
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
 |||||  
 3263 CCTAGGCTACAGAGCTCTGCTAAGAGGCCAGATTTGCAGAGAGAGG 3312  
 |||||  
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
 |||||  
 3313 TAACATACCTGGGTACAGTTTGGGGGCGGCAGCGATGGCTGACGGAG 3362  
 |||||  
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
 |||||  
 3363 GCACGGGAAGAACTGTAGTCCAGATACCGGGCCCAACACACAGCCAAACA 3412  
 |||||  
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
 |||||  
 3413 AGTCAGAGAGTTTGGGCACAGCTGGATTTTGCAGACTGTGGATCCCGG 3462  
 |||||  
 434 lypPheAlaThrLeuAlaProLeuTyrProLeuThrLysGlnLysGly 450  
 |||||  
 3463 GGTTCGACCTTAGCAGCCCCACCTCTACCCGCTAACCAAGAAAGG 3512  
 |||||  
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
 |||||  
 3513 GGATTCCTCGGGCTCTCAGCACCAAGAGGCAATTTGATGCTATCAAAA 3562  
 |||||  
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
 |||||  
 3563 GGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAACCC 3612  
 |||||  
 484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
 |||||  
 3613 TTACCCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAAACC 3662  
 |||||  
 501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
 |||||  
 3663 CAAACCCCTAGGACCATGGAGAGACCTGTGCTACCTGCTCAAGAGAGCT 3712  
 |||||  
 517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534  
 |||||  
 3713 TGATCTGTAGGCAGTGGTTGGCCGTATGCTGAAGGCTATCGCAGCTG 3762  
 |||||  
 534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
 |||||  
 3763 TGGCCATCTGGTCAAGGACGCTGACAAATTCACCTTTGGGACAGAATA 3812  
 |||||  
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
 |||||  
 3813 ACTGTAATAGCCCCCATGTCATPGAGAACATCGTTGCGCAGCCGCCAGA 3862  
 |||||  
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
 |||||  
 3863 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCCTGCTCTCA 3912  
 |||||  
 584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
 |||||

3913 CAGAGAGGTCACCTTCGCTCCACAGCGCTCTCAACCCCTGCCACTCTT 3962  
 |||||  
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
 |||||  
 3963 CTGCGCTGAAGAGACTGATGAACACAGTACTCATGATTGCCATCAACTATT 4012  
 |||||  
 617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
 |||||  
 4013 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4062  
 |||||  
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
 |||||  
 4063 GAGAAGTCTTAACCTGGTTCACTGACGGAAGCAGCTATGTTGGTGAAGGT 4112  
 |||||  
 651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
 |||||  
 4113 AAGAGATGGCTGGGGCGGCAGTGGTGACGGGACCGCACGATCTGGGC 4162  
 |||||  
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
 |||||  
 4163 CAGCAGCCTGCCGGAAGAACTTCAGCGCAAAAGGCTGAGCTCATGGCCC 4212  
 |||||  
 684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
 |||||  
 4213 TCACGCAAGCTTTGGGCTGGCCGGAAGGAATCCATAAACATTTATACG 4262  
 |||||  
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
 |||||  
 4263 GACAGCAGTATGCTTTGGGACTGCACACGTACACGGGGCCATCTATAA 4312  
 |||||  
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
 |||||  
 4313 ACAAGGGGGTGTCTTACCTCAGCAGGGAGGGAATAAAGAACAAAGAGG 4362  
 |||||  
 734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
 |||||  
 4363 AAATTCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4412  
 |||||  
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLys 767  
 |||||  
 4413 ATACACTGCTCGACATCAGAAAGCCAAAGATCTCATATCTAGAGGAA 4462  
 |||||  
 767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784  
 |||||  
 4463 CCAGATGGCTCACCGGGTTGCCAAGCAGCGCAGCCAGGCTGTTAAACCTTC 4512  
 |||||  
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
 |||||  
 4513 TGCCTATAATAGAAACGCCCAAGCCCAAGAACCCAGACGACGACAGTACAC 4562  
 |||||  
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
 |||||  
 4563 CTAGAAGACTGGCAAGAGATAAAAAGATAGACCAGTCTCTGAGACTCC 4612  
 |||||  
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
 |||||  
 4613 GGAGGGGACCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACCAAG 4662  
 |||||  
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
 |||||  
 4663 AAGGTTAGAATAATGTCCACAGATACATCGTCTAACCCACCTAGGAAC 4712  
 |||||  
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
 |||||  
 4713 AACACCTGCAGCAGTTGGTGCAGAACATCCCTTATCATGTTCTGAGGCT 4762  
 |||||  
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
 |||||  
 4763 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATTTGTGCGCCCTGCCAGCTGG 4812  
 |||||  
 884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
 |||||



1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
1249 GGGAGAGCGGTTCGGACCCCTCCCGAGGCCAGGGTAATTTGAAGGT 2198  
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
1299 GGAGGGCAACCACTTGAGTTCCCTGGTTGATACCGGAGCGGAGCATTCAG 2248  
34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
2249 TGCTGCTACACCAATTTAGAAAATTAAGAAAAAATCTGGGTGATG 2298  
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAs 67  
2299 GGTGCCACAGGGCAACGCGCATATCCATGGACTACCCCGAAGACCGTTGA 2348  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84  
2349 CTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGCTCATCCCTGAGTGCC 2398  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
2399 CAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2448  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
2449 TCTTTTGAACAAAGCAAGCAAGAGTGTCTGTGAATAACAAACCCATCAC 2498  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
2499 TGTGTTGACCCCTCAATTAGATGATGAATATCGGACTATATCTCCCAAG 2548  
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
2549 TAAAGCCTGATCAAGATATACAGTCCCTGGTGGAGCAGTTCGCCAAGCC 2598  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
2599 TGGCGAAACCGCAGGAGTGGGTTTGGCAAGCAAGTTCCCCACAGGT 2648  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
2649 TATTTCACTGAAGCCAGTGTACACAGTATCAGTCAGACAGTACCCCT 2698  
184 euSerArgGluAlaArgGluGlyIleThrProHisValGlnArgLeuIle 200  
2699 TGAGTAGAGGCTCGAGAGGAATTTGGCCGATGTTCAAGATTAATC 2748  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
2749 CAACAGGGCATCTAGTTCCTGTCNAATCCCTTGGAAATCTCCCTGCT 2798  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
2799 ACCGGTTAGGAAGCCTGGGACCAATGATATCGACAGTACAGGACTTGA 2848  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
2849 GAGAGGTCAATAAAAGSGTCAGGACATACACCCAGGTCGCCGACCCCT 2898  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpThrValle 267  
2899 TATAACCTCTTGAGCGCCCTCCCGCTTGAACGGAACTGGTACACAGTATT 2948  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
2949 GGACTTAAAGATGCTCTTCTTCCTGAGATTACACCCACCTAGGCAAC 2998  
284 roLeuPheAlaPheGluThrArgAspProGlyThrGlyArgThrGlyGln 300  
2999 CACTTTTGGCTTCGAATGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3048  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317

3049 CTCACCTGGACCGACTGCCCAAGGGTTCAGAACTCCCGACCAATCTT 3098  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3099 TGACGAAGCCCTACACAGGGACCTGGCCAATTCAGGATCCCAACACCTTC 3148  
334 lNValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAlaThr 350  
3149 AGGTGACCCCTCTCCAGTACGTGATGAGCTGCTTCTGGCGGAGCACCC 3198  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
3199 AAACAGSACTGCTTACAGGTAGAGGCACTACTGCTGGAATTTGCTGA 3248  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3249 CTTAGGCTACAGAGCCTCTGCTAAGAAAGGCCAGATTTGCAGAGAGAGG 3298  
384 aThrTrpLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
3299 TAACATACTTGGGTACAGTTTTCGGGGCGGCGAGCATGGCTGACGGAG 3348  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
3349 GCACGGAAGAAACTCTAGTCCAGATACCGGCCCAACACAGCCAAACA 3398  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
3399 AGTGAGAGAGTTTTGGGGACAGCTGGATTTTGACACTGTGGATCCCGG 3448  
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
3449 GGTTCGACCTTAGCAGCCCTCTACCCCTAACCAAGAAAGAGG 3498  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
3499 GGATTCCTCTGGCTCTCAGCAGCAGAGGCAATTTGATGCTATCAAAA 3548  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
3549 GGCCTCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAACCCCT 3598  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
3599 TTACCTCTTATGTGGATGAGCTAGGGAGTAGCCCGAGGAGTTTAACC 3648  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
3649 CAAACCTTAGGACCATGGAGGAGACCTGTTGCTACCTGTCAAGAAAGCT 3698  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
3699 TGATCTGTAGCCAGTGGTTGGCCGCTATGCTGAAGGCTATCGCAGCTG 3748  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
3749 TGGCCATACCTGTCAGGAGCTGACAAATTCACCTTTGGGACAGAAATA 3798  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
3799 ACTGTAATAGCCCGCATGCAATGGGAACATCGTTCCGCGAGCCCCAGA 3848  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
3849 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCTGTCTCTCA 3898  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
3899 CAGAGAGGTCACCTTTCGCTCCACGCGCTCTCAACCTGCCACTCTT 3948  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617

3949 CTGCGCTGAAGAGACTGTGTAACCACTGACTCATGATTGCCATCAACTATT 3998  
617 uileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
3999 GATTGAGGAGACTGGGGTCCGACAGGACCTTACAGACATACCGCTGACTG 4048  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
4049 GAGAAGTGTAACTTGGTTCCTACTGACGGAAGCAGCTATGTGTGGAAGGT 4098  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4099 AAGAGGATGGCTGGGGCGGAGTGGTGACGGGACCCCGACAGTCTGGGC 4148  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4149 CAGCAGCTTGGCGGAAGAACTTACGCCAAAGGCTGAGCTCATGGCCC 4198  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4199 TCACGCAAGCTTGGCGGTGGCGGAAGGAAATCCATAAATTTATACG 4248  
701 aspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
4249 GACAGCAGGTATGCCCTTTGGCAGCTGCACACGTACACGGGGCCATCTATAA 4298  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4299 ACAAGGGGGTGTCTTACCTCAGCAGGGGAGGAAATAAAGAACAAAGAGG 4348  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4349 AAATTTCTAAGCCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATT 4398  
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4399 ATACACTGTCTCTGGACATCAGAAAGCCAAAGATCTCATATCTAGAGGGA 4448  
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAlaAsnLeuL 784  
4449 CCAGATGGCTGACCGGGTTGCCAAGCAGGACGCCAGCTGTAACTTTC 4498  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
4499 TGCCTATAATAAGAACGCCCAAGCCCAAGCCAGACACGACATACACC 4548  
801 LeuGluAspTrpGlnIleLysLysIleAspGlnPheSerGluThrPr 817  
4549 CTAGAAAGCTGCCAAGAGATAAAAGATAGACACGATCTCTGAGACTCC 4598  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
4599 GGAGGGGACCTGCTATACCTATATATGGGAAGGAAATCCTGCCCAAAAG 4648  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
4649 AAGGGTTAGAATATCTCAACAGATACATCGTCTAAACCCACTAGGAAC 4698  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
4699 AAACACCTGCACGACTGGTCAGAACATCCCTTATCATGTTCTCGAGGCT 4748  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
4749 ACCAGAGTGGCTGACTCGGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4798  
884 aLysAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
4799 TTAATGCTAATCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAAGC 4848  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
4849 CACCAGGCGCTCACTGGGAAGTGGACTTCCTAGGAGTAAAGCGGCTAA 4898

917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
4899 ATACGGAACAATAATATCTATTGTTTTTGTAGACACCTTTTCAGGATGG 4948  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
4949 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 4998  
951 ileLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
4999 ATACTGGAGGAATAATTTTCCAGATTTGGATACCTAAGTAATAGGGTC 5048  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5049 AGACAATGTCTCAGCTTTCGTTGCCAGGTAAAGTCAGGACTGGCCAAGA 5098  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5099 TATTGGGGATTGATTGGAACCTGATTTGCATACAGACCCCAAGCTCA 5148  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
5149 GGCAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACCAAT 5198  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
5199 GACCACAGAGACTGGCATTAATGATTGGATGGCTCTCTGCCCTTTGTGC 5248  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
5249 TTTTGGGTGAGGAACACCCCTGGACGTTTGGGCTGACCCCTATGAA 5298  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
5299 TTGCTCTACGGGGGACCCCGCTGGCAGAAATTCCTTTGCACATAG 5348  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
5349 TGCTGATGTCTGCTTTCAGCAGCTTGTCTTAGGCTCAAGGCGCTCG 5398  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
5399 AGTGGGTGAGGACGAGCGTGGAGCAGCTCCGGAGGCGCTACTCAGGA 5448  
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117  
5449 GGAGCTTGCAAGTTCACATCGCTTCCAAAGTTGGAGATTCACTCTATGT 5498  
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134  
5499 TAGACGCCACCGTGCAGGAACCTCGAGACTCGGTGGAGGGACCTTATC 5548  
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150  
5549 TCGTACTTTTGACCAACACACCGCTGTGAAGTCGAGGAATCCCCACC 5598  
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167  
5599 TGGATCCATGCATCCCAACGTTAAGCYGGCGCCACCTCCCGACTCGGGGTG 5648  
1167 pArgAlaGluLysThrCluAsnProLeuLysLeuArgLeuHisArgLeuV 1184  
5649 GAGAGCGGAAAAAGACTGAGAAATCCCTTAAAGCTTCGCTCCATCGCCTGG 5698  
1184 alProTyrSerAsnAsnSerProGlyGln 1194  
5699 TTCTTACTCTAACATAACTCCCGAGGCCAG 5730

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: AAT74884  
seq\_documentation\_block:  
ID AAT74884 standard; cdna; 7892 bp.



XX

AC AAT74884;

XX

DT 09-FEB-1998 (first entry)

XX

DE Miniature swine retrovirus cDNA.

XX

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
KW xenotransplantation; infectious; provirus; organ transplant; donor;  
KW activated virus; PCR; ss.

XX

OS Porcine retrovirus.

XX

FH Key Location/Qualifiers

FT CDS

585..2159

FT

/\*tag= a

FT

mat\_peptide

FT

585..2156

FT

/\*tag= b

FT

/\*note= "putative GAG protein"

FT

CDS

FT

2307..5744

FT

/\*tag= c

FT

mat\_peptide

FT

2307..5741

FT

/\*tag= d

FT

/\*note= "putative POL protein"

FT

CDS

FT

5620..7536

FT

/\*tag= e

FT

mat\_peptide

FT

5620..7533

FT

/\*tag= f

FT

/\*note= "putative ENV protein"

FT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Percent Similarity: 98.992 Ratio: 5.173 Gaps: 1

Percent Identity: 96.893

alignment\_block:

US-09-171-553B-5 x AAT74884

Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

|||||

2160 GGGAGACGGGGTTCGGACCCCTCCCGAGCCAGGGTAACCTTTGAAGT 2209

|||||

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34

|||||

2210 GGAGGGGCAACCACTTGCTGCTGATACCGAGCGAACAACATTGAG 2259

|||||

34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50

|||||

2260 TGCTACTACAGCCATTAGGAAACATAAAGATAAAAAATCTCTGGGTGATG 2309

|||||

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67

|||||

2310 GGTGCCACAGGGCAACAACAGTATCCATGGACTACCCGGAAGAACAGTTGA 2359

|||||

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84

|||||

2360 CTGGGAGTGGGACGGGTAACCCACTCGTTCTGTGTCATACCTGAGTGCC 2409

|||||

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100

|||||

2410 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGAGCACAAATT 2459

|||||

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117

|||||

2460 TCTTTTGAACAAGGGAAACCAAGAGTGTCTGCAAAATAACAAACCTATCAC 2509

|||||

117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnV 134

|||||

2510 TGTGTTGACCCCTCCCAATTAGATGAGCAATATCGACTATCTCTCCCTAG 2559

|||||

134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150

|||||

2560 TAAAGCCTGATCAAAATATACAATTTCTGGTTGGAACAGTTTCCCAAGCC 2609

|||||

151 TrpAlaGluThrAlaGlyMetClyLeuAlaLysGlnValProGlnVa 167

|||||

2610 TGGGCAGAAACCGCAGGGATGGGTTGGCAAGCAAGTTCCCCCAAGT 2659

|||||

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

|||||

2660 TATTCACTGAAGCCAGTGCACACCACTGTCTAGTCAGACAGTACCCT 2709

|||||

184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeu 200

|||||

2710 TGAGTAAAGAAGCTCAAGAGGAATTCGGCGCATGTCCTCAAAAGATTAACT 2759

|||||

201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217

|||||

2760 CAACAGGCGATCTCTAGTTCCTCTCCCAATCTCCCTGGAATCTCTCCCTGCT 2809

|||||

217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234

|||||

2810 ACCGGTTAGAAGGCTGGGACTAATGACTATCGACACAGTACAGGACTTGA 2859

|||||

234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250

|||||

2860 GAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAACCTT 2909

|||||

251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267

|||||

2910 TATAACCTCTGTGCTCTCCCAACCAACCGGAGCTGGTATACAGTATT 2959

|||||

267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284

|||||

|||||

New nucleic acid from porcine retroviruses - used for detecting  
viruses in transplant or other tissue and for assessing risk of  
transmitting infection to graft recipient

Claim 22; Fig 3; 128pp; English.

This cDNA sequence represents a porcine retrovirus from miniature swine  
containing the putative coding regions for viral GAG, POL and ENV  
proteins. This sequence and PCR fragments generated from it  
(see AAT74812-T74882) can be used to screen organs for the presence of  
porcine retroviruses prior to xenotransplantation. Transplantation can  
increase the likelihood of retroviral activation if intact and  
infectious proviruses are present. The porcine retroviral sequence can be  
used to generate probes to determine the level (e.g. copy number) of  
intact (i.e. potentially replicating) porcine provirus sequences in a  
strain of xenograft transplantation donors. It can be used to detect  
mutations, genetic lesions or viral recombinants and also to determine  
the histological localisation of activated retrovirus. Using Polymerase  
Chain Reaction DNA Quantitation (PQO) on blood mononuclear cells  
infectivity titration and susceptibility testing can be performed.  
Ultimately animal donors without intact porcine retroviral sequences or a  
lower copy number of viral elements could be selected.

Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

alignment\_scores:

Quality: 6099.50

Length: 1191

|||||  
2960 GGACTTAAGGATGCTTCTTCCCTCAGATTACACCCCACTAGCCAAAC 3009  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3010 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3059  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3060 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCACTT 3109  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3110 TGACCAAGCCCTACACAGAGACCTGGCCCAACTTCAGGATCAACACCCCTC 3159  
334 InValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAlaThr 350  
|||||  
3160 AGGTGACCCCTCCTCAGTACCTGGATGACCTGCTTCTGGCGGGAGCCACC 3209  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
|||||  
3210 AACAGAGACTGCTTAGAGGCCACGAAGGCACCTACTGCTGGAAATGTCGTA 3259  
367 pLeuGlyTyTrpAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
|||||  
3260 CCTAGGCTACAGAGCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGAGG 3309  
384 alThrTyTrpLeuGlyTyTrpSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
|||||  
3310 TAACATACTTGGGTACAGTTTGGCGGACGGGCAGCGATGGCTGACGGAG 3359  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
|||||  
3360 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3409  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
3410 AATGAGAGAGTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 3459  
434 lyPheAlaThrLeuAlaAlaProLeuTyProLeuThrLysLysGly 450  
|||||  
3460 GGTTCGACACTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGGG 3509  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
3510 GAATTCCTCTGGCTCCTCAGCAGCACCAGAAAGCATTTGATGCTATCAAAA 3559  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
3560 GGCCCTTGCTGAGGGCACCTGCTTGGCCCTCCTGACGTAACCTAACCCCT 3609  
484 heThrLeuTyTrpValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
3610 TPACCCCTTATGAGGATGAGCGTAAAGGAGTAGCCCGGGAGTTTAAACC 3659  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
3660 CAAACCCCTAGGACCATGGAGAGACCTGTCGCTACCTGTCAAAGAAGCT 3709  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534  
|||||  
3710 CGATCTGTAGCCAGTGGTGGCCCATATGCTTGAAGGCTATCGCAGCTG 3759  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
3760 TGCCCATCTGGTCAAGGAGCGCTGACAAATTTGACTTTGGACAGAAATA 3809  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
3810 ACTGTAATAGCCCCCATGTCATTGGAGAACATGCTTTCGCGAGCCCCCAGA 3859  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyTrpGlnSerLeuLeuLeu 584  
|||||  
3860 CCATGATGATGACCAAGCCCGCATGACCCCACTATCAAAAGCCCTGCTTCTCA 3909  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
3910 CAGAGAGGTCACGTTGCTTCCACAGACCGCTCTCAACCTGCCACTCTT 3959  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
3960 CTCCTGAAGAGACTGATGAACCACTGACTCATGATTGTCATCACTATT 4009  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4010 GATTGAGGAGACTGGGGTCGCAAGACCTTCACAGACATACCGCTGACG 4059  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyTrpValValGluGly 650  
|||||  
4060 GAGAACTGCTAACCTGCTTCACTGACGGAAGCAGCTATGTGTGGAAGGT 4109  
651 LysArgMetAlaGlyAlaAlaValAspGlyThrArgThrIleTrpAl 667  
|||||  
4110 AAGAGGATGCTGGCGCGGGTGGTGGACGGGACCCGACGATCTGGGC 4159  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684  
4160 CAGCAGCTGCCGAGAGAACTTCAGACAAAAGGCTGAGCTCATGGCCC 4209  
684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyTrp 700  
4210 TCAGCAAGCTTTGCGGCTGGCGGAAGGAAATCCATAAACTTATACG 4259  
701 AspSerArgTyTrpAlaPheAlaThrAlaHisValHisGlyAlaIleTyLy 717  
4260 GACAGCAGTATGCTTGGCTGCTGACGTACACGTACATGGGGCCATCTATAA 4309  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4310 ACAAGGGGGTGTCTTACCTCAGCAGGGGGAATAAAGAACAAAGAGG 4359  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4360 AAATTCCTAAGCTATTAGAACCCGTACATTTACCAAAAGGCTAGCTATT 4409  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4410 ATACACTGCTCTGGACATCAGAAAGCTTAAGATCTCATATCCAGAGAAA 4459  
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784  
4460 CCAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGCCAGGGTGTAAACCTC 4509  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyTrp 800  
4510 TGCCTATAATAGAAATGCCAAAGCCCAAGAACCCAGACAGACATACACC 4559  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
4560 CTAGAGACTGGCAAGAGATATAAAGATAGACCAGTCTCTGAGACTCC 4609  
817 oGluGlyThrCysTyTrpThrSerTyTrpGlyLysGluIleLeuProHisLysG 834  
4610 GGAAGGGACCTGCTATACCTCAGATGGGAGGAAATCTGCCCCCAAG 4659  
834 luGlyLeuGluTyTrpValGlnGlnIleHisArgLeuThrHisGlyThr 850  
4660 AAGGTTTAGAATATGTCACAGATACATCGTCTAACCCCACTAGGAAC 4709  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyTrpHisValLeuArgLe 867  
4710 AAACACTGCGACAGCTTGGTCAGAACATCCCCCTTATCATGTTCTGAGGCT 4759  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
4760 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATGTGTGCCCTGCCAGCTGG 4809

```

884 alaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
      ::::::::::::::::::::
4810 TTAATGCTAATCTCCAGAACTCCCTCCAGGGAAGAGACTAAGGGCAAGC 4859

901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaIy 917
      ::::::::::::::::::::
4860 CACCCAGGCGCTCCTCGGGAAGTGGACTTCCTAGGTAAAGCGCGCTAA 4909

917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934
      ::::::::::::::::::::
4910 ATACGGAAACAATACTATTGTTTTTGTAGACACCTTTTCAGGATGGG 4959

934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
      ::::::::::::::::::::
4960 TAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGTGGCTAAATAA 5009

951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
      ::::::::::::::::::::
5010 ATACTGGAAGAAATTTTCCAGATTGGAATACCTTAAGTAAATAGGGTC 5059

967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
      ::::::::::::::::::::
5060 AGACAATGTCAGCTTTTGTGCCAGGTAACTCAGGACCTGGCCAGA 5109

984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
      ::::::::::::::::::::
5110 TATTGGGATTGATTGGAACCTGCATTGTGCATACAGACCCCAAGCTCA 5159

1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
      ::::::::::::::::::::
5160 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAAT 5209

1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
      ::::::::::::::::::::
5210 GACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCTGCGCCCTTGTGC 5259

1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
      ::::::::::::::::::::
5260 TTTTGGGTAGGAACACCCCTGCACAGTTTGGCTGACCCCTATGAA 5309

1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067
      ::::::::::::::::::::
5310 TTACTCTACGGGGGACCCCGCCATTTGGTAGAAATTTGCTGTACATAG 5359

1067 rAlaAspValLeuLeuSerGlnProPhePheSerArgLeuLysAlaLeuG 1084
      ::::::::::::::::::::
5360 TGCTGACGTCGCTTTCCAGCCCTTGTCTCTAGGCTCAAGGCACCTTG 5409

1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099
      ::::::::::::::::::::
5410 ACTGGGTGAGACACACGAGCGTGGAGGCAACTCCGGGAGGSCCTACTCAG 5459

1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116
      ::::::::::::::::::::
5460 GGAGGAGACTTCAGATCCCATCGTTTCCAAAGTGGGAGATTCACTCTA 5509

1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133
      ::::::::::::::::::::
5510 CATTAGACCCACCGCTGAGGAAACCTCGAGACTCGGTGGAGGCGCCCTT 5559

1133 yLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149
      ::::::::::::::::::::
5560 ATCTCGTACTTTTGACCAACACACGCGCTGTGAAGTGCAGGAATCTCC 5609

1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGly 1166
      ::::::::::::::::::::
5610 ACCTGGATTCATGATCCACGTTAAACCGGCGCCACCTCCCGATTCGGG 5659

1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183
      ::::::::::::::::::::
5660 GTGGAAACCCGAAAGACTGAATATCCCTTAAAGCTTCGCGCTCCATCGCG 5709

```

```

1183 euValProTyrSerAsnAsnAsn 1190
      ::::::::::::::::::::
5710 TGGTTCTTACTCTGTCAATAC 5732

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF77727
seq_documentation_block:
ID   AAF77727 standard; DNA; 8132 BP.
XX
AC   AAF77727;
XX
DT   23-MAY-2001 (first entry)
XX
DE   Nucleotide sequence of a retrovirus found in miniature swine.
XX
KW   Retrovirus; graft transplantation; xenotransplantation; miniature swine;
XX   ss.
XX
OS   Unidentified.
XX
PN   US6190861-B1.
XX
PD   20-FEB-2001.
XX
PF   13-DEC-1996; 96US-0766528.
XX
PR   14-DEC-1995; 95US-0572645.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Fishman JA;
XX
WPI: 2001-256211/26.
P-PSDB; AAB73285, AAB73286, AAB73287.

Assessing risk of endogenous retroviruses in clinical practice and in
xenotransplantation, comprises using probe sequences derived from swine
or miniature swine retroviral genome -

Claim 1; Fig 3; 127pp; English.

The present invention relates to a method for screening a cell or tissue
for the presence or expression of a retrovirus (RV), comprising
contacting a target nucleic acid from the cell or tissue with a second
nucleic acid from the present invention (e.g. the present sequence or a
fragment thereof). The method is useful for RV detection and to assess
graft transplantation risk. Screening of animals allows the elimination
of donors with active replication of known viruses. Inactive proviruses
can be detected and inactivated, allowing identification and elimination
of potential human pathogens derived from swine in a manner not possible
in the outbred human organ donor population and is important to the
development of human xenotransplantation.

Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment_scores:
Quality: 6099.50 Length: 1191
Ratio: 5.173 Gaps: 1
Percent similarity: 98.992 Percent Identity: 96.893

alignment_block:
US-09-171-553B-5 x AAF77727 ..

Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
      ::::::::::::::::::::
2160 GGGAGACGGGGTTCGGACCCCTCCCGACCCAGGGTAACTTTGAGGT 2209

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34
      ::::::::::::::::::::
2210 GGAGGGCAACCACTTGAGTTCTGTTGTATACCGGAGCGAAACATTCAG 2259

```

34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysSerTrpValMet 50  
|||||  
2260 TGCTACTACAGCCATTAGGAAACTAAAAAGATAAAAAATCCTGGTGATG 2309  
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrThrArgArgThrValAs 67  
|||||  
2310 GGTGCGACAGGGCAACACAGATATCCATGGACTACCCGGAAGACAGTTGA 2359  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84  
|||||  
2360 CTTGGGAGTGGGACGGGTAACCCACTCGTTCTGTGTATACCTGAGTGCC 2409  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
|||||  
2410 CAGCACCCCTCTTAGGTAGACATTATTGACCAAGATGGGAGCACAAATT 2459  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
|||||  
2460 TCTTTTGAACAGGGAACCAAGAGTGTCTGCAAAATACAAACATATCAC 2509  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
|||||  
2510 TGTGTTGACCCCTCCNAATTAGATGACGAATATCGACTATCTCTCCCTAG 2559  
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
|||||  
2560 TAAAGCCTGATCAAAATATACAATTCTGGTTGGAACAGTTTCCCCAAAGCC 2609  
151 TrpAlaGluThrAlaGlyMetClyLeuAlaLysGlnValProProGlnVa 167  
|||||  
2610 TGGGCGAAGAACCGCAGGAGTGGTTTGGCAAGCAAGTTCCCCACCAAGT 2659  
167 IleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
|||||  
2660 TATTCACTGAAGGCCAGTGCACACACAGTGTCACTGACGACAGTACCCT 2709  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
|||||  
2710 TGAGTAAAGAAAGCTCAAGAAAGAAATTCGGCGGCATGTCCAAGATTAAATC 2759  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
|||||  
2760 CAACAGGCATCTAGTTCCTCCCAATCTCCCTGGAAATACCTCCCTGCT 2809  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
2810 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCATCGACAGACATTGA 2859  
234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
2860 GAGAGGTCAATAAAGGGGTGAGGATATACACCCCAACAGTCCCGAACCCCT 2909  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValIle 267  
|||||  
2910 TATAACCTCTCTGTGCTCTCCACCCCAACGAGCTGGTATACAGATT 2959  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
|||||  
2960 GGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAAC 3009  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3010 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCCGGGCGAG 3059  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3060 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCACTT 3109  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3110 TGACGAAGCCCTTACACAGAGACCTGGCCAACTTCAGGATCCAAACACCCCTC 3159  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
334 InValThrLeuLeuGlnTyrValAspLeuLeuAlaGlyAlaThr 350  
|||||  
3160 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3209  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
|||||  
3210 AACAGGACTGCTTAGAAGGCACGAGGCACTACTGCTGGAATTTGCTGA 3259  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluV 384  
|||||  
3260 CCTAGGCTACAGAGCCTGCTAAGAAGGCCAGATTTCAGGAGAGAGG 3309  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
|||||  
3310 TAAACATACTTGGGTACAGTTTTCGGGACGCGAGCTGGTGGTGGAG 3359  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
|||||  
3360 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3409  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
3410 AATGAGAGAGTTTTGGGACACAGCTGGATTTCGACAGCTGTGATCCCGG 3459  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
3460 GGTTCGGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAGAGG 3509  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
3510 GAATTCCTCTGGGCTCTTGAGCACACAGAGGCATTTCGATATCAAAA 3559  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
3560 GGCCCTGCTGAGCGCACCTGCTTGCCCTCCTGACCTAACTAAACCCCT 3609  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
3610 TTACCTCTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACC 3659  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
3660 CAAACCTAGGACCATGGAGAACCTGTGCGCTACCTGTCAAAAGAGCT 3709  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaV 534  
|||||  
3710 CGATCTCTGTAGCCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCTG 3759  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
3760 TGCCCATACTGGTCAAGGACGCTGACAAATTCGACTTTGGGACAGAAATA 3809  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
3810 ACTGTAATAGCCCCCATGCTGATTGGAGAACATCGTTTCGCGAGCCCCCAGA 3859  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
3860 CCGATGGATGACCAACGCGCATGACCCACTATCAAGGCTGCTCTCTCA 3909  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
3910 CAGAGAGGTCAGTTCGCTCCACAGCCGCTCTCAACCTGCCACTCTT 3959  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
3960 CTCCTGAACAGACTGATGAACCCAGTCACTCATGATTGCCATCAACTATT 4009  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4010 GATTGAGGAGACTGGGGTCGCAAGGACCTTACAGACATACCCGCTGACTG 4059  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650

||||| 4060 GAGAAAGTCTAACCTGGTTCCTACTGACGGAACAGCATATGTGTGGGAAGGT 4109  
||||| 651 LysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAl 667  
||||| 4110 AAGAGGATGGTGGGGCGCGGTGGTGGACGGGACCCGACGATCTGGGC 4159  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4160 CAGCAGCCTGCCGGAAGAACTTCAGCACAAAAGGCTGAGCTCATGGCCC 4209  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
||||| 4210 TCAGCAGAGCTTTCGGCGTGGCCGGAAGGAAATCCATAAATTTATACG 4259  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
||||| 4260 GACAGCAGGTATGCTTTGGCAGTGCACACGTACATGGGGCCATCTATAA 4309  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4310 ACAAGGGGGTGTCTTACCTCAGCAGGAGGGGAAATAAGAACAAAGAGG 4359  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4360 AAATCTTAAGGCTATTAGAAGCCGTACATTTACCAAAAAGGCTAGCTATT 4409  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
4410 ATACACTGTCTGGACATCAGAAAGCTTAAAGATCTCATATCCAGAGGAAA 4459  
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784  
4460 CCAGATGGCTGACCGGGTTGCCAAGCAGGAGCCAGGGTGTAAACCTTC 4509  
784 euProIleIleGluThrProLysAlaProGluProArgGlnIleTyrThr 800  
4510 TGCCTATAATAGAAATGCCCAAGCCCGCCAGAACCCAGACAGCATACACC 4559  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
4560 CTAGAACACTGGCNAAGAGATAAANAAGATAGACCAGTTCTCTGAGACTCC 4609  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
4610 GGAAGGGACCTGTCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAG 4659  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
4660 AAGGTTTAGAATATGTCACACATACATCGTCTAACCCACCTAGGAACCT 4709  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
4710 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCT 4759  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
4760 ACCAGGAGTGGCTGACTCGGTGGTCAAAACATTTGTGCGCCCTGCCAGCTGG 4809  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
4810 TTAATGCTAATCTTCAGAAATGCCCTCCAGGGAAGAGACTTAAGGGAAGC 4859  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
4860 CACCCAGGCGCTCACTGGGAAGTGGACTTCACCTGAGGTAAAGCCGCTAA 4909  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpv 934  
4910 ATACGGAACAAATACTATTGTTTTTGTAGACACCTTTTCAGGATGGG 4959  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||

4960 TAGAGGGTTATCTCTACTAGAAAAGAGACTTCAACCGTGGTGGCTAAAAA 5009  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967  
||||| 5010 ATACTGGAAGAAATTTTCCAGATTTGGAATACCTTAAGGTAATAGGGTC 5059  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5060 AGACAATGGTCCAGCTTTTGTGCCCGAGGTAAAGTCAGGACCTGGCCAGA 5109  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5110 TATTGGGGATTGATTGGAACCTGATTGTGCATACAGACCCCAAGCTCA 5159  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
5160 GGACAGGTAGAGAGGATGAATAGAACCATTTAAGAGACCCCTTACTAAAT 5209  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
5210 GACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCTCTGCCCTTTGTGC 5259  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
5260 TTTTATAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5309  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
5310 TTACTCTACGGGGACCCCCCATTTGGTAGAAATTTGCTGTACATAG 5359  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
5360 TGTCTGACGTGCTCTTCCAGCCTTTGTCTCTAGGCTCAAGGCACTTG 5409  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
5410 ACTGGGTGAGACAACAGCGCTGGAGCAACTCCGGGAGGCTTACTCAGGA 5459  
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116  
5460 GGAGGAGACTTCAGATCCCATCGTTTCCCAAGTGGGAGATTCAGTCTA 5509  
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpTlysGlyProT 1133  
5510 CGTTAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGCCCTT 5559  
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
5560 ATCTCGTACTTTTGACCACACACAGCGCTGTGAAAAGTCGAAGAAATCTCC 5609  
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166  
5610 ACCTGGATCCATGCATCCACGTTAAACCGGGCCACCTCCCGATTCGGG 5659  
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
5660 GTGGAAGCCGAAAGACTGAAAATCCCTTAAGCTTCGCTCCCTCATCGG 5709  
1183 euValProTyrSerAsnAsnAsn 1190  
5710 TGGTTCTCTACTCTGTCAATAAC 5732

seq\_name: /SIDS2/gcgdata/geneseq/NA2001.DAT: AAC67020

seq\_documentation\_block:

ID AAC67020 standard; DNA: 4402 BP.

XX AAC67020;

XX AC

XX DT 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 20.

XX

KW Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX W0200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000MO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX PA Federspiel MJ;

XX PI WPT; 2001-032041/04.

XX DR Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA

PT encoding extracellular proteins of the agent into donor cells, such as

PT swine cells -

XX Claim 16; Page 105-106; 144pp; English.

XX The present invention provides a method to prevent the transmission of

CC infectious agents during xenotransplantation. This involves introducing

CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the

CC transplant recipient.

XX Sequence 4402 BP; 1259 A; 1085 C; 1111 G; 947 T; 0 other;

# alignment\_scores:

Quality: 6095.50 Length: 1192

Ratio: 5.174 Gaps: 1

Percent Similarity: 98.826 Percent Identity: 96.980

## alignment\_block:

US-09-171-553b-5 x AAC67020 ..

Align seg 1/1 to: AAC67020 from: 1 to: 4402

1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

224 GGAACACGGGGTTCGGACCCCTCCCGGAGCCAGGGTAACCTTGAAGGT 273

17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisServ 34

274 GGAGGGCAACACAGTTGAGTTCTGTTGATACCGGAGCGAAACATTTCAG 323

34 aLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50

324 TGCTACTACAGCCATTAGGAAACCTAAAGATATAAAATCCCTGGGTGATG 373

51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValas 67

374 GGTGCCACAGGCGAACACAGTATCCATCGACTCCCGAAGAACAGTTGA 423

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84

424 CTGGGAGTGGGAGGGTAACCCACTCGCTTCTGCTCATACCTGAGTGCC 473

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyValaGlnIle 100

474 CAGACCCCTCTAGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 523

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117

524 TCTTTTGAACAGGAAACACAGAGTGTCTGCAAAATAACAAACCVATCAC 573

117 rValLeuThrLeuGlnLeuAspGluTyrArgLeuLysSerProGlnV 134

574 TGTGTTGACCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 623

134 allYsProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150

624 TAAAGCTGATCAAAATATACAAATCTGGTTGGAACAGTTTCCCAAGCC 673

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167

674 TGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCACAAGT 723

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

724 TATTCAACTGAAGGCCAGTGCCACACAGTGTCTGTCAGTCAGACAGTACCCT 773

184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200

774 TGAGTAAGAGCTCAAGAGGAATTCGGCCGCATGTCCAAAGATTAATC 823

201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217

824 CAACAGGGCATCTAGTTCTGTCCAATCTCCCTGGAATACTCCCTGTCT 873

217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234

874 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCAGTACAGGACTTGA 923

234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250

924 GAGAGTCAATAAACGGTGCAGATATACACCAACAGTCCCGAACCT 973

251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267

974 TATAACCTCTGTCTCTCCACCCCAACGAGGAGTGTGTATACAGTATT 1023

267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284

1024 GGACTTAAAGGATGCCCTTTCTGCTGAGATTACACCCCACTAGGCAAC 1073

284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300

1074 CACTTTTTCCTTCGATGGAGATCCAGTACGGGAGAACCGGCGAG 1123

301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317

1124 CTCACCTGGACCCGACTGCCCCAAGGGTTCAGAACTCCCGGACCATCTT 1173

317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334

1174 TGCGAAGCCCTACACAGAGACCTTGGCCAACTTCAGGATCCAAACCCCTC 1223

334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350

1224 AGGTGACCTCTCCAGTACGTGATGACCTGCTCTCTGGCGGAGCACCC 1273

351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367

1274 AAACAGACTGCTTAGAAGGCACGAGGCACTACTGCTGGAATGTCTGA 1323

367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384

1324 CCTAGGCTACAGACCTCTCTAGAGGCCAGATTTCGAGGAGAGAGG 1373

384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400

1374 TAACATACTTGGGGTACAGTTTGGCGACGCGGACGATGGCTGACGGAG 1423

401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417

1424 GCAGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCAGCAACA 1473

417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434

1474 AGTGAGAGAGCTTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCGG 1523  
 434 lyPheAlaThrLeuAlaAlaProLeuThrProLeuThrLysGluLysGly 450  
 1524 GTTTGGGAGCTTAGCAGACCCCACTTACCCACTAACCAAGAAAGG 1573  
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
 1574 GAATTCCTCTGGGCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 1623  
 467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysPro 484  
 1624 GGCCCTCTGAGCGCACTCTCTGGCCCTCCCTGAGTAACATAAACCCCT 1673  
 484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500  
 1674 TTACCTTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGGAGTTTAAACC 1723  
 501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
 1724 CAAACTTAGGACCATGGAGGAGACCTGTTGCCTACCTGTCAAGAAGCT 1773  
 517 wAspProValAlaSerClyTrpProValCysLeuLysAlaIleAlaLav 534  
 1774 CGATCCTGTAGCCAGTGGTTGGCCCTGATGCCCTGAAGGCTATCCGAGCTG 1823  
 534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
 1824 TGGCCATCTGTTCAAGGACCGTGCACAAATTGACTTTGGGACAGAATAA 1873  
 551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
 1874 ACTGTAATAGCCCCCATGCTTGGAGAACATCGTTTCGGCAGCCCCCAGA 1923  
 567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
 1924 CCGATGGATGACCAACGCCCGCATGACCCCACTATCAAGCCCTGCTTCTCA 1973  
 584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600  
 1974 CAGAGAGGTTCAGTTCGCTCCACAGCCGCTCTCAACCCCTGCCACTCTT 2023  
 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
 2024 CTGCCTGAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATT 2073  
 617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
 2074 GATTGAGGAGACTGGGGTCGCGAAGGACCTTACAGACATACCCGCTGACTG 2123  
 634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
 2124 GAGAAGTGTAACTGTTCACTGACGGAAGGACGCTATGTGTGGGAAGGT 2173  
 651 LysArgMetaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
 2174 AAGAGGATGCTGGGCGGGGTGGTGGACGGGACCCGACGATCTGGGC 2223  
 667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetaLAl 684  
 2224 CAGCAGCCTGCGGGAAGGAACTTCAGCACAAAAGGCTGAGCTCATGGCC 2273  
 684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700  
 2274 TCACGAAGCTTTGGCGCTGGCGAAGGAAATCCATAAACATTTATACA 2323  
 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
 2324 GACAGCAGGTATGCTTTTGGACTGTCACAGCTACACGGGGCCATCTATAA 2373  
 717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
 2374 GCAAAGGGGCTTGTCTTACCTCAGCAGGGAGGAAATAAAGAACAAAGAG 2423

734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
 2424 AAATCTTAAGCCTATTAGAAAGCCTTACATTTGCCAAAAGGCTAGCTATT 2473  
 751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
 2474 ATACACTGTCTGGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 2523  
 767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
 2524 CCAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGGGGTGTAACCTTC 2573  
 784 euProIleIleGluThrProLysAlaProGluProArgArgGlnThrThr 800  
 2574 TGCCCTATGATAGAAACACCCAAAGCCCAAGACCCGAGCAGGACGACACC 2623  
 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
 2624 CTAGAAGACTGGCAAGAGATAAAAAGATAGACCAGTTCTCTGACACTCC 2673  
 817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
 2674 GGAAGGAGCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAG 2723  
 834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
 2724 AAGGTTTGAATATATGTCCAACAGATACATCGTCTAACCCACCTAGGAAC 2773  
 851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
 2774 AAACACCTGCACAGTGTGTCAGAACATCTCCTTATCATGTCTGAGGCT 2823  
 867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
 2824 ACCAGAGTGGCTGATTCGGTGGTCAACACTGTGTGCCCTGCCAGCTGG 2873  
 884 alAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900  
 2874 TTAATGCTTAATCCTTCCAGATACCTCCAGAAAGAGACTAAGGGGAAGC 2923  
 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
 2924 CACCAGGCGCTCACTGGCAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 2973  
 917 styGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
 2974 ATACGGAAACAAATATCTATTGGTTTTTTGTAGACACCTTTTCAGGATGG 3023  
 934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
 3024 TAGAGGCTTATCTCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAGAAA 3073  
 951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
 3074 ATACTGGAGGAAATTTTCCAAGATTTGGAATACCTAAGGTAATAGGTC 3123  
 967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
 3124 AGACAATGGTCCAGCTTTCGTTGCCAGGTAAAGTCAGGAGCTGGCCAAAG 3173  
 984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
 3174 TATTTGGGATTTGATTGGAACTGTCATTGTGCATACAGACCCCAAGCTCA 3223  
 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
 3224 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAAT 3273  
 1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
 3274 GACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCCTCGCCCTTGTGTC 3323

1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
 3324 TTTTGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 3373  
 1051 LeuLeuTyrGlyProProLeuAlaGluLeuAlaPheAlaHisSe 1067  
 3374 TTACTCTACGGGGACCCCTTGTGAGAAATGCTTCCGTCACATAG 3423  
 1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
 3424 TGCTGACGTGCTTTCCTCCAGCCTTGTCTTAGGCTCAAGGCACCTG 3473  
 1084 LuTrpValArgGlnAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
 3474 AGTGGGTGAGACAACGAGCGTGGAGGCAACTCCGGGAGGCTACTCAGGA 3523  
 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116  
 3524 GGAGGAGACTTGCAGATCCACATCGTTCCTCAAGTGGGAGATTCAGTCTA 3573  
 1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133  
 3574 CGTAGAGCCACCGTCGAGAAACCTCGAGACTCGGTGGAGGGCCCTT 3623  
 1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
 3624 ATCAGTACTTTTACCACACCAACGCGTGTGAAAGTCGAAGGAATCTCC 3673  
 1150 ThrTrpIleHisAlaSerHisValLysProAlaProProAspSerG1 1166  
 3674 ACCTGGATTCATGATCCACGTTAAGCGCGCCACCTCCCGATTCCGG 3723  
 1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
 3724 GTGGAACCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCCATCGCG 3773  
 1183 euValProTyrSerAsnAsnSer 1191  
 3774 TGGTTCCTTACTCTGTCAATAACTCC 3799

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC67019

seq\_documentation\_block:

ID AAC67019 standard; DNA; 7362 BP.

XX AAC67019;

XX 27-MAR-2001 (first entry)

XX PERV env protein coding sequence SEQ ID NO: 19.

XX xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 99US-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian

XX transplant recipients, by introducing recombinant DNA comprising DNA

XX encoding extracellular proteins of the agent into donor cells, such as

XX swine cells.

XX

PS Claim 16; Page 101-104; 144pp; English.

XX

CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.

XX

SQ Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

alignment\_scores:

Quality: 6095.50 Length: 1192

Ratio: 5.174 Gaps: 1

Percent Similarity: 98.826 Percent Identity: 96.980

alignment\_block:

US-09-171-553B-5 x AAC67019 ..

Align seg 1/1 to: AAC67019 from: 1 to: 7362

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
 2823 GGAAGACGGGTTTCGGACCCCTCCCGAGCCAGGGTAACCTTGAAGGT 2872  
 17 lGluGlyGlnProValIgluPheLeuValAspThrGlyAlaGluHisSerV 34  
 2873 GGAGGGCAACCACTTTCCTGTTGATACCGGAGCGAAACATTCAG 2922  
 34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
 2923 TGTACTACAGCCATTAGGAAACATAAAGATAAAATAATCCCTGGGTGATG 2972  
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67  
 2973 GGTGCCACAGGGCAACACAGATATCCATCGACTACCCGAGAACAGTTGA 3022  
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
 3023 CTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGTGTCATACCTGAGTGCC 3072  
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
 3073 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGACACAAAT 3122  
 101 SerPheGluGlnGlyArgProGluValSerValAsnLysProIleTh 117  
 3123 TCTTTTGACACAGGGAACACAGAGTGTCTGCAATAACAAACCTATCAC 3172  
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
 3173 TGTGTTGACCCCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 3222  
 134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
 3223 TAAAGCCTGATCAAAATATACAAATCTGTTGGACACAGTTTCCCAAGCC 3272  
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
 3273 TGGGAGAAACCCGAGGATGGTTTGGCAAGCAAGTTCCCCCAAGT 3322  
 167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
 3323 TATTCAACTGAAGGCCAGTCCACACAGTGTGTCAGTCAGACAGTACCOCT 3372  
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
 3373 TGAGTAAAGAAAGCTCAAGAGGAATTCGGCCGCGATGTCCAAAGATTATC 3422  
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
 3423 CAACAGGGCATCTCTCTGTGCCAATCTCCCTGGGAATACTCCCTGCT 3472



217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
3473 ACCGGTTAGAAAGCCTGGGACTATGACTATCCAGCAGTACAGACTTGA 3522  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
3523 GAGAGGTCAATAAACGGGTGAGGATATACACCAACAGTCCCGAACCT 3572  
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpThrValle 267  
|||||  
3573 TATAACCTCTGTGTCTCTCCCAACCGGAGGTGGTATACAGTAT 3622  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnp 284  
|||||  
3623 GGACTTAAGAGTGCCTTTTCTGCTGAGATTACACCCCACTAGCCAAC 3672  
284 roLeupheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3673 CACTTTTTCGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 3722  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3723 CTACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3772  
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
|||||  
3773 TGACGAAGCCCTACACAGAGACCTGCCCAACTTCAGGATCCCAACACCTC 3822  
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
|||||  
3823 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCACC 3872  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367  
|||||  
3873 AAACAGACTGCTTAGAAGGCACGACGAGGCACTACTGCTGGAATGTCTGA 3922  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
|||||  
3923 CCTAGGCTACAGAGCCTCTGCTAAGAAGCCCAAGATTTGCAGGAGAGAG 3972  
384 alThrTrpLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
|||||  
3973 TACATACTTGGGTACAGTTTTCGGGACGGCAGCGATGGCTGACGGAG 4022  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
|||||  
4023 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCCCAACA 4072  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
4073 AGTGAGAGAGTTTGTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGG 4122  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGlyLysGly 450  
|||||  
4123 GGTTCGACCTTAGCAGCCCACTCTACCCACTAACCAAGAAAGAGG 4172  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
|||||  
4173 GAATCTCTCTGGCTCCTGAGCACAGAGGCAATTTGATGTATCAAAAA 4222  
467 salAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
4223 GGCCCTGTGAGCGACCTGCTCTGGCCCTCCCTGACCTAACTAACCCCT 4272  
484 heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500  
|||||  
4273 TTACCTTTATGTGATGAGCGTAAAGGAGTAGCCCGGGGATTTTAACC 4322  
501 GlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLe 517  
|||||  
4323 CAAACTCTAGGNCCATGAGGAGACCTGTTCCTACCTACCTGTCAAGAGACT 4372  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIav 534  
|||||  
4373 CATCTCTGTAGCCAGTGGTGGCCGTATGCTGAAGGCTATCGCAGCTG 4422  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
4423 TGGCCATACTGCTCAAGGACGCTGACAAATGACTTTGGGACAGAAATA 4472  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
4473 ACTGTAATAGCCCCCATCGTTGGAGAACATCGTTCCGACGCCCCAGA 4522  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
|||||  
4523 CCGATGGATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTTCTCA 4572  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
4573 CAGAGAGGTCACTGCTCCACCGCGCTCTCAACCCCTGCCACTCTT 4622  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
|||||  
4623 CTGCTCTGAAGAGACTGATGAACAGTACTCATGATTGCCATCAACTATT 4672  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
|||||  
4673 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 4722  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4723 GAGAAGTGTAACTGTTTCACTGACGGAAGCAGCTATGTGTGGAGAGT 4772  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4773 AAGAGATGGCTGGGGCGGGTGGTGACGGGACCCCGCAGCATCTGGGC 4822  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
|||||  
4823 CAGCAGCCTGCCGAGAGAACTTCAGCACAAAGGCTGAGCTCATGGCCC 4872  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4873 TCACGCAAGCTTTCGGCTGCGGAGGAAATCCATAAACATTTATACA 4922  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4923 GACAGCAGGTATGCTTTGGACTGCACTACACGGGGCCCATCTATAA 4972  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
|||||  
4973 GCAAGGGGGTGTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 5022  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
|||||  
5023 AAATTTCAAGCCTATTAGAAGCCTTACATTTGCCAAAAAGGCTAGCTATT 5072  
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
|||||  
5073 ATACACTGTCTGGACATCAGAAAGCCCAAGATCCCATATCCAGAGGGAA 5122  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
|||||  
5123 CCAGATGCTGACCGGGTTGCCAAGCAGCAGCCCGGCTGTTAACTTC 5172  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
5173 TGCCTATGATAGAAACACCCCAAGCCCGAGAACCCGACGACAGTACACC 5222  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5223 CTAGAACACTGGCAAGAGATAAAAAAGATAGACAGTTCTCTGAGACTCC 5272  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834

||||| 5273 GGAAGGGACCTGCCTATACCTCAGATGGGAGGAAATCCTGCCCCACAAAG 5322  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
||||| 5323 AAGGGTTAGAAATATGTCACACAGATACATCGTCTAACCCACCTAGGAAT 5372  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
||||| 5373 AAACACCTGCAGCAGTGGTTCAGAACATCTCTTATCATGTCTTCAGGCT 5422  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
||||| 5423 ACCAGGAGTGGCTGATTCCGGTGGTCAAAACACTGTGTGCCCTGCCAGCTGG 5472  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
||||| 5473 TTAATGCTTAATCTTCCAGATACCTCCAGGAAAGAGACTAAGGGGAGC 5522  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
||||| 5523 CACCCAGGCGCTCACTGGGAAGTGGACTTCACCTGAGGTAAAGCCGCTAA 5572  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
||||| 5573 ATACGGAACAATAATATCTATTGTTGTTTGTAGACACCTTTTTCAGGATGGG 5622  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
||||| 5623 TAGAGCGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAA 5672  
951 IleLeuGluGluIlePhePheArgPheGlyIleProLysValIleGlyse 967  
||||| 5673 ATACTGGAGGAAATTTTCCAAAGATTGGAAATACCTAAGGTAAATAGGTC 5722  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
||||| 5723 AGACAATGTCAGCTTTCGTTGGCCAGGTAAAGTCAGGAGTGGCCCAAGA 5772  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
||||| 5773 TATTGGGATGATTGGAACCTGCATTGTCATACAGACCCCAAGCTCA 5822  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
||||| 5823 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGAGACCTTACTAAAT 5872  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
||||| 5873 GACCGGGAGACTGGCGTTAATGATTGATAGTCTCTCTGCTGCTTGTAC 5922  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
||||| 5923 TTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5972  
1051 LeuLeuTyrGlyIleProProLeuAlaGluIleAlaPheAlaHisSe 1067  
||||| 5973 TTAATCTACGGGGGACCCCCCATTTGGTAGAAATTTGCTTCGTCATAG 6022  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
||||| 6023 TGCTGACGTGCTGCTTTCACAGCCTTGTCTCTAGGCTCAAGGCATTTG 6072  
1084 luTrpValArgGlnArgAlaTrpLysLysLeuArgGluAlaTyrSer... 1099  
||||| 6073 AGTGGGTGAGACAACAGCGGTGGAGCAACTCCGGGAGGCGCTTACTCAGGA 6122  
1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTy 1116  
||||| 6123 GGAGAGACTTCAGATCCCATCGTTTCCAAAGTGGGAGATTCAGTCTA 6172  
1116 rValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProT 1133  
|||||

6173 CGTTAGACGCCACCGTGCAGAGAAACCTCGAGACTCGGTGGAAGGCCCTT 6222  
1133 yrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIlePro 1149  
||||| 6223 ATCACCTACTTTTGACCAACACCGCTGTGAAAGTCGAAGAAATCTCC 6272  
1150 ThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerG 1166  
||||| 6273 ACCTGGATCATGCATCCACCGTTAAGCGGGGCCACCTCCCGATTCCGG 6322  
1166 yTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgL 1183  
||||| 6323 GTGGAAGCGGAAAGAGCTGAAATCCCTTAAGCTTCGCTCCATCCGCG 6372  
1183 euValProTyrSerAsnAsnSer 1191  
||||| 6373 TGGTTCCTTACTCTGTCAATAACTCC 6398

seq\_name: /SID82/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67023

seq\_documentation\_block:

ID AAC67023 standard; DNA; 7873 BP.

AC AAC67023;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

PN WO200071726-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14296.

XX 24-MAY-1999; 9905-0135631.

XX (MAYO-) MAYO MEDICAL VENTURES.

XX Federspiel MJ;

XX WPI; 2001-032041/04.

XX Inhibiting or preventing infectious agent transmission in mammalian  
transplant recipients, by introducing recombinant DNA comprising DNA  
encoding extracellular proteins of the agent into donor cells, such as  
swine cells -

PS Claim 16; Page 112-115; 144pp; English.

XX The present invention provides a method to prevent the transmission of  
infectious agents during xenotransplantation. This involves introducing  
to donor swine cells a recombinant DNA encoding a peptide fragment from  
the infectious agent, and then introducing these cells into the  
transplant recipient.

XX Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

alignment\_scores:

Quality: 5858.00 Length: 1194

Ratio: 5.103 Gaps: 4

Percent Similarity: 96.147 Percent Identity: 94.054

alignment\_block:

US-09-171-553B-5 x AAC67023 ..

Align seg 1/1 to: AAC67023 from: 1 to: 7873

```

1  GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
1637  GGAAGACGGGGTTCGGACCCCTCCCGAGCCCGGTAACCTTTGAAGGT 1686
17  lGluGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
1687  GGAGGGCAACCAAGTTGAGTCTCTGGTTGATACCGGAGCGAAACATTGAG 1736
34  alLeuLeuGlnProLeuGlyLysLeuLysGlyLysLysSerTrpValMet 50
1737  TGCTACTACAGCCATTAGGAATACTAAAGATAAAAAATCCTGGGTGATG 1786
51  GlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAs 67
1787  GGTGCCACAGGGCAACAACAGTAGTATCCATGGACTTACCAGGAACAGTTGA 1836
67  pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
1837  CTTGGGAGTGGGACGGGTAAACCACCTCGTTTCTGGTCATACCTGAGTGCC 1886
84  roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
1887  CAGCACCCCTCTTAGTAGAGACTTATTCACCAAGATGGGACCAACAATT 1936
101  SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
1937  TCTTTTGAACAAGGGAACACAGAAGTGTCTGCAAAATAACAAACCTATCAC 1986
117  rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
1987  TGTGTTGACCCCTCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 2036
134  alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
2037  TAAAGCCTGATCAAAATATACAAATCTGTTGGAAACAGTTTCCCAAGCC 2086
151  TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
2087  TGGGCAGAAACCGCAGGGATGGTTTGGCAAAAGCAAGTTCCCCCAACAGT 2136
167  lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
2137  TATTCAACTGGAAGGCGAGTGCACACCACTGTCAGTCAGACAGTACCCT 2186
184  euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
2187  TGAGTAAGAAGCTCAAGAGGAATTCGCCCGCATGTCCAAAGATTATC 2236
201  GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2237  CAACAGGGCATCCTAGTTCTCTCCATCTCCCTTGGAACTACCTCCCTGCT 2286
217  uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2287  ACCGGTTAGAAAGCCTGGGACTTAATGACTATGACCACGATACAGGACTTGA 2336
234  rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2337  GAGAGTCAATAAAGCGGTGAGGATATACACCCCAACAGTCCCGAACCTT 2386
251  TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267
2387  TATAACCTCTTGTGTGCTCTCCACCCCAACGAGCTGGTATACAGTATT 2436
267  uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
2437  GGACTTAAGGATGCTCTTTCTGCTCAGATTACACCCCACTAGCCCAAC 2486
284  roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
2487  CACTTTTTCCTTCGAATGGAGATCCAGGTACGGGGAAGAACCGGCAG 2536
301  LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
2537  CTCACCTGACCCGACTCCCAAGGTTCAAGAACTCCCGACCATCTT 2586
317  eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
2587  TGACGAAGCCCTACACAGAGACTGGCCAACTTCAGGATCCAACACCTC 2636
334  lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
2637  AGGTGACCCCTCTCCAGTACGTGGATGACCTTCTGGCGGAGCCACC 2686
351  LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
2687  AAACAGGACTGCTTAGAAGACGACGAGCACTACTGCTGGAATGCTCTGA 2736
367  pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
2737  CCTAGGCTACAGAGCTCTGCTAAGAGGCCAGATTTGCAGAGAGAGG 2786
384  alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400
2787  TAACATACCTTGGGGTACAGTTTGGGGGACGGGACGATGGCTGACGGAG 2836
401  AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
2837  GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCCCAACA 2886
417  nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
2887  AGTGAGAGAGTTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCG 2936
434  lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
2937  GGTTTGCGACCTTAGCAGCCCACTCTACCCACTAACCAAGAAAGAAAGG 2986
451  GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
2987  GAATTCCTCGGCTCCTGACACACAGAGCATTTGATGCTATCAAAA 3036
467  sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3037  GGCCTGTGTCAGCGCCTGCTCTGGCCCTCCCTGAGCTAACTAACCT 3086
484  heThrLeuTyrValAspGluArgLysGlyValAlaAlaArgGlyValLeuThr 500
3087  TTACCTTTATGTGGATGAGGTAAAGGAGTAGCCCGGGGAGTTTAAAC 3136
501  GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
3137  CAAACTCTAGACCATGGAGAGACCTTGGCTTACCTGTCAAGAAAGCT 3186
517  uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaIaV 534
3187  CGATCTGTAGCCAGTGGTTGGCCGCTATGCTGAAGGCTATCGCAGCTG 3236
534  alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
3237  TGGCCATACCTGGTCAAGGACGCTGACAAATTGACTTTGGGACAGAAATA 3286
551  ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
3287  ACTGTAATAGCCCCCATGGTGGAGAACATCGTTCCGCGAGCCCCCAGA 3336
567  pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584
3337  CCGATGATGACCAACACCGCCGATGACCCACTATCAAGGCTGCTTCTCA 3386
584  hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600
3387  CAGAGAGGTACAGTTCGCTCCACCGCCGCTCTCAACCTGCTGCTCTT 3436
601  LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617

```

```

3437 CTGCCTGAAGAGACTGATGATCAACCAAGTCACTCATGATTGCCATCAACTATT 3486
617 uileGluThrGlyValArgLysAspLeuThrAspLeuProLeuThrG 634
3487 GATTGAGAGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTG 3536
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrrValValGluGly 650
3537 GAGAAGTGTAACTGGTTCACTCAACGGAAGCAGCTATGTAGTGAAGGT 3586
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
3587 AAGAGGATGGCTGGGGGGGGTGGTGGACGGGACCGCAGATCTCGGC 3636
667 aserSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
3637 CAGCAGCCTGCCGAGGAAGTTCAGCACAAAGGCTGAGCTCATGGCC 3686
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrr 700
3687 TCAGCAAGCTTTCGGCTGGCCGGAAGGAAATCCATAAATTTATACA 3736
701 AspSerArgTyrrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrr 717
3737 GACAGCAGGTATGCTTTCGACTGCAACAGTACACGGGGCCATCTATAA 3786
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGlu 734
3787 GCAAAGGGGGTGTCTACCTCAGCAGGGAGGGAATTAAGAACAAAGAGG 3836
734 luileLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
3837 AAATCTAAGCCTATTAGAAGCTTACATTTCGCAAAAGGCTAGCTATT 3886
751 lleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
3887 ATACACTGCTCGACATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 3936
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValaValaLeu 784
3937 CCAGATGGCTGACCGGGTTGCCAAGCAGCAGCAGCCAGGGTGTAAACCTC 3986
784 euProIleIleGluThrProLysAlaProGluProArgArgGluTyrrThr 800
3987 TGCTTATGATGAACACCCCAAGGCCAGAACCCGACGACAGATACACC 4036
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817
4037 CTAGAAGACTGGCAAGACATAAAGATAGACACAGTCTCTGAGACTCC 4086
817 oGluGlyThrCysTyrrThrSerTyrrGlyLysGluIleLeuProHisLys 834
4087 GGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCAAAAG 4136
834 luGlyLeuGluTyrrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
4137 AAGGCTTGAATATGTCACAGATACATCTCTTAACCCACCTAGGAAC 4186
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrrHisValLeuArgLe 867
4187 AAACACCTGCACAGTGTGGTCAAGACATCTCTTATCATGTCTGAGGCT 4236
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 883
4237 ACCAGAGTGGCTGATTGGTGGTCAACACTGTGTGGCTGCCAGCTGG 4286
884 ValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySe 900
4287 GTA.....AAGCCGGCTA 4289
900 rHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaL 917
4290 .....AAGCCGGCTA 4299

```

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67032
seq_documentation_block:
ID AAC67032 standard; DNA; 6076 BP.

```



401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGI 417  
|||||  
3958 GCACGGAAGAAACGTGTAGTCCAGATACCGGCCCAACACAGCCAAACA 4007  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
|||||  
4008 AGTGAGAGAATTTTGGGGCAGCGTGGATTTTGCAGACTGTGGATCCCGG 4057  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
|||||  
4058 GGTTCGCCACCTTAGACGCCCACTCTACCCCACTAACCAAGAAAAGGG 4107  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
4108 GAATTCCTCGTGGCTCTGAGCACCAGAGGCATTTGATGCTATCAAAA 4157  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
|||||  
4158 GGCCCTGTGAGCGCACCTGTCTGGCCCTCCCTGACGTAACCTAAACCT 4207  
484 heThrLeuTyrValAspGluArgLysGlyValAlaAtrGlyValIleuThr 500  
|||||  
4208 TTACCTTTATGTGATGAGCGTAGGGAGTAGCCCGGGAGTTTAAACC 4257  
501 GlnThrLeuGlyProTrpArgProValAlaLayrLeuSerLysLysLe 517  
4258 CAAACTCTAGGACCATGAGGAGACCTCTTGCTACCTGTCAAGAAGCT 4307  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
4308 CGATCTCTGAGCCAGTGTGGCCGCTATGCTGAAGGCTATCGCAGCTG 4357  
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
4358 TGGCCATACGTGTGAAGGACGCTGACAAATTGACTTTGGGACAGNAYATA 4407  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
4408 ACTGTAATAGCCCCCATCGTGTGGAGAACATCGTTCGGCAGCCCCAGA 4457  
567 parGTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
4458 CCGATGGATGACCAACGCCCGCATGCCCACTATCAAAAGCCTGCTCTCA 4507  
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600  
4508 CAGAGAGGTCACGTTCGCTCCACCGCGCTCTCAACCTGCCACTCTT 4557  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
4558 CTGCCTGAAGACATGATGAACCACTGATCATGATTGCCATCAACTATT 4607  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4608 GATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACTG 4657  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
4658 GAGAAAGTGTAACTGGTTCACTGACGAGAGCAGCTATGTGTGGAGGT 4707  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
4708 AAGAGGATGGCTGGGGCGCGGTGGTGGACGGGACCCGACGATCTGGGC 4757  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4758 CAGCAOCCCTGCCGAGAGGAATTCACACAAAGGCTGAGCTCATGGCCC 4807  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
4808 TCACGCAAGCTTTGGCGCTGGCCGAGGAGGAAATCCATAAACATTTATACA 4857

701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
|||||  
4858 GACAGCAGGTATGCCCTTTCGACTGCACACGTACACGGGGCCATCTATATA 4907  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4908 GCAAGGGGGTTCCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAGG 4957  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4958 AAATTCCTAAGCCCTATTAGAAGCCCTTACATTTGCCAAAAAGGCTAGCTATT 5007  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
5008 ATACACTGCTCGCATCAGAAAGCCAAAGATCCCATATCCAGAGGAA 5057  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
5058 CCAGATGGCTGACCGGTTGCCAAGCAGGAGCCAGGGGTGTTAACCTTC 5107  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
5108 TGCCTATGATAGAAACACCCAAAGCCCAAGACCCGCGAGCACAGTACACC 5157  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
5158 CTAGAAGACTGGCAAGAGATAAAAAAGATAGACCACTCTCTGAGACTCC 5207  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
5208 GGAAGGACCTGCTATACCTCAGATGGAGGAAATCCTGCCCCACAAAG 5257  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
5258 AAGGTTTAGAATATGTCCAACAGATACATCGTCTAACCACCTTAGGAAC 5307  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
5308 AAACACCTGCAGCAGTTGGTCAGAACATCTCTTATCATGTTCTGAGGCT 5357  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884  
5358 ACCAGGAGTGGCTGATTTCGGTGTCAAACACTGTGTGCCCTGCCAGCTGG 5407  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
5408 TTAATGTCTAATCTTCCAGAATACCTCCAGAAAGAGACTAAGGGGAAGC 5457  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
5458 CACCCAGCGCTCACTGGGAAGTGGACTTCACCTGAGGTAAAGCGGCTAA 5507  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
5508 ATACGAAACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGG 5557  
934 alGluAlaTyrProThrLysGluThrSerThrValValAlaLysLys 950  
5558 TAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCCGTGGTGGCTAAGAAA 5607  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
5608 ATACTGGAGGAAATTTTCCAGATTGGAAATACCTAAGGTAATAGGGTC 5657  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5658 AGACAAATGGTCCAGCTTTTCGTTGCCAGGTAAAGTCAGGGACTGGCCCAAG 5707  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5708 TATTGGGATTAATTGGAACTGCTTGTGATACACACCCCAAGCTCA 5757  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017

```

|||||
5758 GGACAGCTAGAGGATGAATAGAACATTAAAGAGACCCCTTACTAAATT 5807
|||||
1017 wThrThrCluThrClyleAsnAspTrpMetAlaLeuLeuProPheVal 1034
|||||
5808 GACCGCGAGAGCTGGCGTTAATGATTGGATAGTCTCTCCCTTTGTGC 5857
|||||
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
|||||
5858 TTTTATGGGTAGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907
|||||
1051 LeuLeuTyrGlyGlyProProLeuAlaGluAlaPheAlaHis 1067
|||||
5908 TTACTCTACGGGAGACCCGCCATTGGTAGAAATTGCTTCCGTACATAG 5957
|||||
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeu 1084
|||||
5958 TGCTGAGTCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCATTG 6007
|||||
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer 1099
|||||
6008 AGTGGGTGAGACAAACGAGCGTGGAGCAACTCCGGGAGGCCTACTCAGGA 6057
|||||
1100 GlyGlyAspLeuGlnVal 1105
|||||
6058 GGAGGAGACTTGCAGATC 6075

```

seq\_name: /SID82/gcgdata/geneseq/geneseq/NA1997.DAT: AAT74811

seq\_documentation\_block:

ID AAT74811 standard; cDNA; 8060 BP.

AC AAT74811;

DT 11-FEB-1998 (first entry)

DE Porcine retrovirus Tsukuba-1 cDNA.

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
 KW xenotransplantation; infectious; provirus; organ transplant; donor;  
 KW activated virus; Tsukuba-1; PCR; ss.

OS Porcine retrovirus.

```

FH Key Location/Qualifiers
FT CDS 86..2002
FT FT /*tag= a
FT FT /note= "putative ENV protein"
FT CDS 3112..4686
FT FT /*tag= b
FT FT /note= "putative GAG protein"
FT CDS 4871..8060
FT FT /*tag= c
FT FT /note= "putative POL protein (partial)"

```

PN W09721836-A1.

XX 19-JUN-1997.

XX 13-DEC-1996; 96WO-US19680.

XX 14-DEC-1995; 95US-0572645.

XX (GEHO ) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 1997-332804/30.

XX New nucleic acid from porcine retroviruses - used for detecting  
 PT viruses in transplant or other tissue and for assessing risk of  
 PT transmitting infection to graft recipient

XX

PS Claim 1; Fig 1; 128pp; English.

XX This sequence represents the purified porcine retroviral cDNA  
 CC sequence of Tsukuba-1 and contains the putative coding regions for viral  
 CC proteins GAG, POL and ENV. This sequence and PCR fragments generated  
 CC from the sequence (see AAT74812-T74882) could be used to screen organs  
 CC for porcine retroviruses prior to xenotransplantation. Transplantation  
 CC can increase the likelihood of retroviral activation if intact and  
 CC infectious proviruses are present. The porcine retroviral sequence can be  
 CC used to generate probes to determine the level (e.g. copy number) of  
 CC intact (i.e. potentially replicating) porcine provirus sequences in a  
 CC strain of xenograft transplantation donors. It can be used to detect  
 CC mutations, genetic lesions or viral recombinants and to determine the  
 CC histological localisation of activated retrovirus. Using Polymerase Chain  
 CC Reaction DNA Quantitation (Pdq) on blood mononuclear cells, infectivity  
 CC titration and susceptibility testing can be performed. Ultimately animal  
 CC donors without intact porcine retroviral sequences or with a lower copy  
 CC number of viral elements could be selected.

XX SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment\_scores:

Quality: 5637.50 Length: 1127

Ratio: 5.061 Gaps: 3

Percent Similarity: 98.846 Percent Identity: 96.628

alignment\_block:

US-09-171-553B-5 x AAT74811

Align seg 1/1 to: AAT74811 from: 1 to: 8060

```

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
|||||
4687 GGGAGACGGGGTTCGGACCCCTCCCGAGCCAGGGTAACCTTTCAAGGT 4736

17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
|||||
4737 GGAGGGGCAACAGTTGAGTTCCTGGTTGATACCGGACGGAACATTGAG 4786

34 alLeuLeuGlnProLeuGlyLysLeuLysGlyLysSerTrpValMet 50
|||||
4787 TGCTACTACAGCCATTAGGAAACTAAAGATAAAAAATCTCGGTGATG 4836

51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
|||||
4837 GGTGC.ACAGGGCAACACAGTATCCATGGACTACCCGGAAG.ACAGTTGA 4884

67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84
|||||
4885 CTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGC 4934

84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
|||||
4935 CAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 4984

101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117
|||||
4985 TCTTTTGAACAAAGGGAACACAGAGTGTCTGCAATTAACAAACCTATCAC 5034

117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
|||||
5035 TGTGTTGACCCCTCCCAATTAGATGACGAATATCGACTATATCTCTCCCTAG 5084

134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
|||||
5085 TAAAGCCTGATCAAAATATACAAATCTGGTTGGAAACAGTTTCCCAAGCC 5134

151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167
|||||
5135 TGGCAGAAACCGGAGGATGGGTTTGGCAAGCAAGTTCCCCCAAGT 5184

167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184

```

||||| 5185 TATTCAACTGAAGGCCAGTGCACACAGTGTCAAGTCAGACAGTACCCCT 5234  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuLe 200  
5235 TCAGTAAAGAAGCTCAAGAAGGAATTGGCGCATGTGCCAAGATTAAATC 5284  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
5285 CAACAGGCATCCTAGTTCCTGCTCAATCTCCCTGGAAATACCTCCCTGCT 5334  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
5335 ACCGGTTAGAAAGCTGGGACTAATGACTATCGACACAGTACAGGACTGA 5384  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
5385 GAGAGGTCAATAAAGCGGTGAGGATATACACCAACAGTCCCGAACCT 5434  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
5435 TATAACCTCTTGTGTCTCTCCACCCCAACGAGCTGGTATACAGTATT 5484  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
5485 GGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCAAAC 5534  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
5535 CACTTTTGGCTTGGAAATGGAGATCCAGGTACGGGAAGAACCGGCAG 5584  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
5585 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 5634  
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
5635 TGACGAAGCCCTTACACAGAGACCTGGCCAACTTCAGGATCCAAACCCCTC 5684  
334 InValThrLeuLeuGlnTrpValAspLeuLeuAlaGlyAlaThr 350  
5685 AGGTGACCCCTCTCCAGTACGTGGTGTGATGACCTGCTTGGCGGGAGCCACC 5734  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
5735 AAACAGGACTGCTTGAAGGCACGAAGGCACTACTGCTGGAATTGCTGA 5784  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
5785 CCTAGGCTACAGAGCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGAGG 5834  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400  
5835 TAACATACTGGGTACAGTTTACGGGAGCGGCAGGATGCTGACGGAG 5884  
401 AlaArgLysLysThrValValGlnIleProAlaProThrAlaLysGI 417  
5885 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 5934  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
5935 AATGAGAGAGTATTTGGGACAGCTGGATTTTCAGACTGTGGATCCCGG 5984  
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
5985 GGTTCGACCTTAGCAGCCCACTCTACCCGCTAAACCAAGAAAAAGGG 6034  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
6035 GAATTCCTCGGCTCCTGAGCAGCAGAGGCAATTTGATGCTATCAAAAA 6084  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484

6085 GGCCCTGCTGAGGCGACCTGCTCTGGCCCTCCTGACGTAACTAAACCT 6134  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
6135 TTACCCCTTATGTGATGAGCGTAAGGAGTAGCCGGGGAGTTTAAAC 6184  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
6185 CAAACCTAGGACCATGGAGAAGACCTGTCGCTACCTGTCAAAGAAGCT 6234  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534  
6235 CGATCTGTAGCCAGTGGTGGCCCATATGCCTGAAGGCTATCGCAGCTG 6284  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnII 550  
6285 TGCCCATACTGGTCAAGGACGCTGACAAATGACTTTGGGACAAGAATAT 6334  
550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProA 567  
6335 AACTGTAAATAGCCCCCATGCTTGGAGAACAATCGTTGGCAGCCCCAG 6384  
567 spArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 583  
6385 ACCGATGGATGACCAACGCCCGCATGCCACATATCAAAGCTGCTTCTC 6434  
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600  
6435 ACAGAGGGGTACAGTTGCTGCCCAACACCGCTCTCAACCTGCCACTCT 6484  
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617  
6485 TCTGCTGTAGAGACTGATCAACCACTGACTCATGATTGCCATCAACTAT 6534  
617 eutleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633  
6535 TGATTGAGGAGACTGGGTCCCGAAGACCTTACACACATACCGCTGACT 6584  
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluI 650  
6585 GGAGAAGTGTAACTGTTCACTGACGGAAGCAGCTATGTGTGTGAAGG 6634  
650 yLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpA 667  
6635 TAAGAGATGGCTGGGGCGGGTGGTGGACGGGCCCGCAGCATCTGGG 6684  
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683  
6685 CCAGCAGCTGCCGGGAGGAACCTTCAGCAAAAAGGCTGAGCTCATGGCC 6734  
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleYrTh 700  
6735 CTCACGCAAGCTTTGGGCTGGCCGAAGGAAATCCATAAACATTTATAC 6784  
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleYrL 717  
6785 GGACAGCAGGTATGCCCTTTCGCACTGCACACAGTACATGGGCGCATCTATA 6834  
717 ysGlnArgGlyLeuLeuThrSerAlaGlyArgGluLleLysAsnLysGlu 733  
6835 AACAAAGGGGTTGCTTACCTCAGCAGGAGGAAATAAAGAAACAAAGAG 6884  
734 GluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaI 750  
6885 GAAATTCAGCCCTATTAGAACGCGTATATTACCAAAAAGGCTAGCTAT 6934  
750 eileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767  
6935 TATACACTGCTGGACATCAGAAAGCTAAGATCTCATATCCAGAGAA 6984  
767 snGlnMetAlaAspArgValAlaLysGlnAlaLeuAlaGlnAlaValAsnLeu 783  
6985 ACCAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGGGGTGTTAACCTT 7034



784 LeuProIleLeuGluThrProLysAlaProGluProArgArgGlnTyrTh 800  
|||||  
7035 CTGCTATAATAGAAATGCCAAAGCCCAAGACCCAGACAGTACAC 7084  
800 rLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrP 817  
|||||  
7085 CCTAGAGACTGGCAAGAGATAAAGATAGACCA. TTCTCTGAGACTC 7133  
817 roGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLys 833  
|||||  
7134 CGGAAGGACCTGTATACCTCAGATGGAGGAATCCTGCCCAAAA 7183  
834 GluGlyLeuGluThrValGlnGln. IleHisArgLeuThrHisLeuGlyT 850  
|||||  
7184 GAAGGGTTAGAAATATGTCCAACAAGATACATCGTCTAACCCACCTAGCAA 7233  
850 hrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArg 866  
|||||  
7234 CTAAACACCTGCACGAGTGTGTGAGAATCCCTTATCATGTTCTGAGG 7283  
867 LeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLe 883  
|||||  
7284 CTACAGGAGTGGCTGACTCGTGTCAACATGTGTGCTGCCAGCT 7333  
883 uAlaAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlyS 900  
|||||  
7334 GGTAAATGCTAATCTTCCAGATGCTCCAGGGAAGAGACTAAGGGGAA 7383  
900 erHisProGlyValAlaHisTrpGluValAspPheThrGluValLysProAla 916  
|||||  
7384 GCCACCCAGGGCTTCACGTGGAGTGGACTTCACCTGAGGTAAAGCCGGCT 7433  
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933  
|||||  
7434 AAATATGGAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGGATG 7483  
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950  
|||||  
7484 GGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTAGCTAAA 7533  
950 ysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966  
|||||  
7534 AAATCTGGAAGAAATTTTCCAGATTTTGAATACCTAAGGTAAATAGGG 7583  
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983  
|||||  
7584 TCAGACAATGGTCCAGCTTTTGTGCCAGTAAAGTCAGGAGCTGCCCAA 7633  
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSers 1000  
|||||  
7634 GATATTGGGGATTGATTGGAACCTGCAATTGTCATACAGACCCCAAGCT 7683  
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016  
|||||  
7684 CAGCAGGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTAAA 7733  
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033  
|||||  
7734 TTGACCGCGAGACTGGCGTTATGATTTGATAGCTCTCTGCCCTTTGT 7783  
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050  
|||||  
7784 GCTTTTATAGGTTAGGAACACCCCTGGACAGTTTGGCTTGACCCCTATG 7833  
1050 lLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHis 1066  
|||||  
7834 AATTACTCTACGGGGACCCCCCATTTGTTGATAAATTCCTCTGTACAT 7883  
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083  
|||||  
7884 AGTGCTGATGTGCTGCTTTCCAGACGCTTTGTTCTCTAGGCTCAAGGCAT 7933

1083 uGluTrpValArgGlnAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099  
|||||  
7934 TGAGTGGTGCAGACAAGAGCGTGGAGCAACTCCGGAGGCTACTCAG 7983  
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115  
|||||  
7984 GAGGAGGAGACTTGCAGATCCCATCATCGTTTCCAGTGGGAGATTGAGTC 8033  
1116 TyrValArgArgHisArgAlaGlyAsn 1124  
|||||  
8034 TAGCTTAGACGCCACCGTCAGGAAAC 8060

seq\_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77725

seq\_documentation\_block:

ID AAF77725 standard; cDNA; 8060 BP.

XX AAF77725;

DT 23-MAY-2001 (first entry)

XX Tsukuba-1 cDNA.

XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.  
XX Porcine retrovirus.  
XX OS US6190861-B1.  
XX PN 20-FEB-2001.  
XX PF 13-DEC-1996; 96US-0766528.  
XX PR 14-DEC-1995; 95US-0572645.  
XX PA (GEHO) GEN HOSPITAL CORP.  
XX PI Fishman JA;  
XX WPI; 2001-256211/26.

Assessing risk of endogenous retroviruses in clinical practice and in xenotransplantation, comprises using probe sequences derived from swine or miniature swine retroviral genome -

XX Claim 1; Fig 1; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue for the presence or expression of a retrovirus (RV), comprising contacting a target nucleic acid from the cell or tissue with a second nucleic acid from the present invention (e.g. the present sequence or a fragment thereof). The method is useful for RV detection and to assess graft transplantation risk. Screening of animals allows the elimination of donors with active replication of known viruses. Inactive proviruses can be detected and inactivated, allowing identification and elimination of potential human pathogens derived from swine in a manner not possible in the outbred human organ donor population and is important to the development of human xenotransplantation.

XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment\_scores:

Quality: 5637 50 Length: 1127

Ratio: 5.061 Gaps: 3

Percent Similarity: 98.846 Percent Identity: 96.628

alignment\_block:

US-09-171-553B-5 x AAF77725

Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17

||||| 4687 GGGAGACGGGTTCGGACCCCTCCCGAGCCAGGGTAACTTTGAAGGT 4736  
17 lGluGlnProValGluPheLeuValAspThrGlyAlaGluHiserv 34  
||||| 4737 GGAGGGCAACAGTTGAGTTCCCTGGTTGATACCGGAGCGAAACATTTCAG 4786  
34 aLeuLeuGlnProLeuGlyLysLeuGlyLysLysSerTyrValMet 50  
||||| 4787 TGCTACTACGCCATTAGGAACACTAAAGATAAAAAATCTCGGTGATG 4836  
51 GlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAs 67  
||||| 4837 GGTGC ACAGGCAACACATGATCCATGGACTACCCGGAAG ACAGTTGA 4884  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysp 84  
||||| 4885 CTGGGAGTGGAGCGGTAACCCACTCGTTCTGGTATACCTGAGTGCC 4934  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
||||| 4935 CAGCACCCCTCTTAGGTAGACATTATGACCAAGATGGGAGCACAAATT 4984  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
||||| 4985 TCCTTTGAACAGGGAACACAGAGTGCTGCAAAATACAAACCTATCAC 5034  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
||||| 5035 TGTGTTGACCTCCCAATTAGATGAGCAATATCGACTACTCTCCCTAG 5084  
134 aLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
||||| 5085 TAAAGCCTGATCAAAATATACAAATTCCTGGTTGGAACAGTTTCCCAAGCC 5134  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
||||| 5135 TGGGAGAAACCGCAGGAGTGGTTTGCAAGCAAGTTCCCCACCAAGT 5184  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
||||| 5185 TATTCACCTGAGGCGCAGTCCACACCAAGTGTACGTACAGACAGTACCCCT 5234  
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200  
||||| 5235 TGAGTAAAGAGCTCAAGAAAGGAATTCGCGCGCATGTCCAAAGATTAAATC 5284  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
||||| 5285 CAACAGGGCATCCTAGTTCTCTGTCCTCAATCTCCTGGAAATATCTCCCTGCT 5334  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
||||| 5335 ACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCACTACAGGACTTGA 5384  
234 rGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
||||| 5385 GAGAGGTCAATAAAGCGGTGCGAGGATATACACCAACAGTCCCGAACCCCT 5434  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLe 267  
||||| 5435 TATAACCTCTGTGTGCTCTCCACCCCAACCGGAGCTGGTATACAGTATT 5484  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
||||| 5485 GGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAAC 5534  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
||||| 5535 CACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG 5584  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||

5585 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCCGACCATCTT 5634  
317 eaSpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
||||| 5635 TGACGAAGCCCTACACAGAGACTGGCCAACTTCAGGATCCAAACACCTC 5684  
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
||||| 5685 AGGTGACCTCTCCAGTACGTGGATGACCTGCTCTGGGGGAGCCACC 5734  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAs 367  
||||| 5735 AAACAGGACTGCTTAGAAGCGACGAAGGCACACTACTCTGGAATGTCTGA 5784  
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluV 384  
||||| 5785 CCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGG 5834  
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
||||| 5835 TAAACATACTGGGGTACAGTTTACGGGAGCGGCGAGTGGCTGACGGAG 5884  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
||||| 5885 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCCAACA 5934  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
||||| 5935 AATGAGAGAGTTTTGGGACAGCTGGATTTTCAGACACTGTGGATCCCG 5984  
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
||||| 5985 GGTTCGACCTTAGCAGCCCACTCTACCGCTAACCAAGAAAGAAAGG 6034  
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysly 467  
||||| 6035 GAATTCCTCGGGCTCTGAGCACAGAGGCAATTTGATGCTATCAAAA 6084  
467 saLaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484  
||||| 6085 GGCCCTCTGAGCGCACCTGCTGGCCCTCCCTGACGTAACTAAACCCCT 6134  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
||||| 6135 TTACCTTTATGTGATGAGCGTAGGAGTAGCGCGGAGGAGTTTAAACC 6184  
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517  
||||| 6185 CAACCCCTAGGACCATGGAGAAGACCTGTGCGCTACTCTGTCAAGAAAGCT 6234  
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534  
||||| 6235 CGATCTCTAGCCAGTGGTGGCCCATATGCCCTGAAGGCTATCGCAGCTG 6284  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGln AsnIle 550  
||||| 6285 TGGCCATCTGGTCAAGGACGCTGCACAAATTGACTTTGGACACAGATAT 6334  
550 eThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProA 567  
||||| 6335 AACTGTAAATAGCCCCCATGCATTGGAGAACATCGTTCCGAGCCGCCAG 6384  
567 spArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 583  
||||| 6385 ACCGATGATGACCAACGCGCGATGACCCACTATCAAGCCCTGCTTCTC 6434  
584 ThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLe 600  
||||| 6435 ACAGAGAGGTGACGTTGCTTCCCAACACGCTCTCAACCCCTGCCACTCT 6484  
600 uLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuL 617  
||||| 6485 TCTGCCTGAAGAGACTGATGAACCAAGTAGTGCATGATTGCCATCAACTAT 6534

```
617 euileGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThr 633
|||||
6535 TGATTGAGGAGACTGGGTCGCGAAGACCTTACAGACATACCGCTGACT 6584
634 GlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluG 650
|||||
6585 GGAGAAGTGTAACCTGTTCACTGACGGAAGCAGCTATGTGTGGAAGG 6634
650 YLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpA 667
|||||
6635 TAGAGAGTGGCTGGGCGCGGTGGGACGGAGCCCGCACCATCTGGG 6684
667 laSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 683
|||||
6685 CCAGCAGCCTGCGGAGGAACCTCAGCACAAAGGCTGAGCICATGGCC 6734
684 LeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrTh 700
|||||
6735 CTCACGCAAGCTTTGGGGCTGGCGGAAGGGAATCCATAAACATTTATAC 6784
700 rAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrL 717
|||||
6785 GGCAGCAGGTATGCTTTCGACTGCACACGTACATGGGGCCATCTATA 6834
717 ysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGlu 733
|||||
6835 AACAAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAG 6884
734 GluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaI 750
|||||
6885 GAAATCTAAGCCTATTAGAAGCGTACATTTACCAAAAAGGCTAGCTAT 6934
750 eileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyA 767
|||||
6935 TATACACTGTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAA 6984
767 snGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeu 783
|||||
6985 ACCAGATGGCTGACCGGGTGGCAAGCAGCAGCCCGAGGGTTAACTTT 7034
784 LeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrTh 800
|||||
7035 CTGCCTATATAGAAATGCTCAAAAGCCCAAGACCCAGACAGCAGTACAC 7084
800 rLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrP 817
|||||
7085 CCTAGAAGACTGGCAAGAGATAAAAAGATAGACCA.TTCTCTGAGACTC 7133
817 roGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLys 833
|||||
7134 CGGAAGGAGCTGCTATACCTCAGATGGGAAGGAATCTCGCCCAACAA 7183
834 GluGlyLeuGluTyrValGlnGlnL.IleHisArgLeuThrHisLeuGlyT 850
|||||
7184 GAAGGGTTAGATATATGTCACAAAGATACATGCTCTAACCCACCTAGGAA 7233
850 hrLysHisLeuGlnGlnLeuValIatqThrSerProTyrHisValLeuArg 866
|||||
7234 CTAAACACCTGGCAGCAGTGGTCAGAACATCCCTTATCATGTTCTGAGG 7283
867 LeuProGlyValAlaAspSerValValLysHisCysValProCysGlnLe 883
|||||
7284 CTACAGAGAGTGGCTGACTCGGTGGTCAACATTTGTGTCCTGCCAGCT 7333
883 uValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlys 900
|||||
7334 GGTAAATGTATATCTTCCAGATGCTCCAGGAAGAGACTAAGGGGAA 7383
900 erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAla 916
|||||
7384 GCCACCCAGCGCTCAGTGGGAAGTGGACTTCACTGAGGTAAAGCGCGCT 7433
```

```
917 LysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTr 933
|||||
7434 AAATATGGAAACAATATACCTATTGTTTGTAGACACCTTTTCAGGATG 7483
933 pValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysL 950
|||||
7484 GGTAGAGGGTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTAGCTAAA 7533
950 ysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGly 966
|||||
7534 AAATACTGGAGAAAATTTTTCAGGATTGGAAATACCTAAGGTAATAGG 7583
967 SerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLy 983
|||||
7584 TCAGACAATGGTCCAGCTTTTGTGCCCAGGTAAGTCAGGGACTGGCCAA 7633
983 sIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSers 1000
|||||
7634 GATATTGGGATTTGATTGAAACTGCATTGTGCATACAGACCCCAAGCT 7683
1000 erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLys 1016
|||||
7684 CAGGACAGGTAGAGAGATGAATAGAACCATTTAAAGAGACCTTACTAAA 7733
1017 LeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheVa 1033
|||||
7734 TTGACCGGGAGACTGGCGTTAATGATTGGATAGCTCTCTCGCCCTTTGT 7783
1033 lLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrG 1050
|||||
7784 GCTTTTATAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATG 7833
1050 luLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHis 1066
|||||
7834 AATTACTACGGGGACCCCCCATTTGTAGAAATGCTCTCTGTACAT 7883
1067 SerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLe 1083
|||||
7884 AGTGTGATGTGCTGCTTTCCAGGCTTTGTTCTAGGCTCAAGGCACT 7933
1083 uGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer. 1099
|||||
7934 TGAGTGGGTGAGACAACAGCGGTGGAGGCAACTCCGGGAGGCTTACTCAG 7983
1100 ..GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerVal 1115
|||||
7984 GAGGAGGAGACTTGCAGATCCCATGCTTTTCCAAGTGGGAGATTCAGTC 8033
1116 TyrValArgArgHisArgAlaGlyAsn 1124
|||||
8034 TACGTTAGACCCACCGTGCAGGAAC 8060
seq_name: /STDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC67021
seq_documentation_block:
ID AAC67021 standard; DNA; 6076 BP.
XX
AC AAC67021;
XX
XX 27-MAR-2001 (first entry)
XX
DE PERV env protein coding sequence SEQ ID NO: 21.
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
OS Porcine endogenous retrovirus.
XX
PN W0200071726-A1.
XX
XX 30-NOV-2000.
XX
XX 24-MAY-2000; 2000WO-US14296.
XX
XX
```

PR 24-MAY-1999; 99US-0135631.  
 PA (MAYO-) MAYO MEDICAL VENTURES.  
 PI Federspiel MJ;  
 XX WPI; 2001-032041/04.  
 DR Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 XX Claim 16; Page 107-109; 144pp; English.  
 XX  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 XX Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;

## alignment\_scores:

Quality: 5635.50 Length: 1106  
 Ratio: 5.156 Gaps: 1  
 Percent Similarity: 98.825 Percent Identity: 97.016

## alignment\_block:

US-09-171-553B-5 x AAC67021 ..

Align seg 1/1 to: AAC67021 from: 1 to: 6076

1 GlyArgGlySerAspProGluProGluProArgValThrLeuLysVa 17  
 2758 GGAAGACGGGGTTCGGAGCCCTCCCGAGCCAGGGTAACCTTGAAGGT 2807  
 17 lclGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
 2808 GGAGGGCAACAGCTGGGTTCTCGTTGTATACCGGAGCGAACAATTCAG 2857  
 34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
 2858 TGCTACTACGCCATTAGGAAACTAAAGATAAAATAATCCTGGGTGATG 2907  
 51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67  
 2908 GTGCCACAGGGCAACAACAGATATCCATGGACTACCCGAGAACAGTTGA 2957  
 67 pLeuGlyValGlyArgValThrHisSerPheLeuValileProGluCysp 84  
 2958 CTGGAGTGGGACGGGTAAACCCACTCGTTCTGTATACCTGAGTGCC 3007  
 84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
 3008 CAGCACCCCTCTTAGTAGACTATTATGACCAAGATGGGAGCACAAAT 3057  
 101 SerPheGlnGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
 3058 TCTTTTGAACAGGGAACCAAGAGTGTCTCAATACCAACCTATACAC 3107  
 117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
 3108 TGTGTTGACCTCCCAATTAGATGACGAATATCGACTATACCTCCCTAG 3157  
 134 alLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
 3158 TAAAGCCTGATCAAAATATACAAATCTGGTTGGGAACAGTTTCCCCAGGCC 3207  
 151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnVa 167  
 3208 TGGCGAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCCAACAGT 3257

167 ltleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
 3258 TATTCAACTGAAGGCGAGTGCACACACAGTGTCAAGTCAAGTACCCCT 3307  
 184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuile 200  
 3308 TAGTAAGAAGCTCAAGAGGAATTCGGCGCATGTCCAAAGATTAATC 3357  
 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeu 217  
 3358 CACAGGGCATCTAGTCTCTCCAAATCTCCCTGGAATACTCCCTGCT 3407  
 217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
 3408 ACCGGTTAGAAAGCTTGGGACTAATGACTATGACGAGTACAGGACTTGA 3457  
 234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
 3458 GAGAGGTCAATAAAGGGTGCAGGATATACACCAACAGTCCCGAACCT 3507  
 251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLe 267  
 3508 TATAACCTCTTGTGCTCTCCACCCCAACGAGCTGTGTATACAGTATT 3557  
 267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
 3558 GGACTTAAAGGATGCCCTTTTCTGCTGAGATTACACCCACTACGCCAAC 3607  
 284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
 3608 CACTTTTTCCTTCGAATGGAGATCCAGGATACGGGAAGAACCCGGCAG 3657  
 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
 3658 CTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGGACCATCTT 3707  
 317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
 3708 TGACGAAGCCCTACACAGAGACTTGCCCAACTTCAGGATCCCAACACCTC 3757  
 334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350  
 3758 AGGTGACCTCTCCAGTACGTGGGATGACCTGCTTCTGGCGGAGCCACC 3807  
 351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367  
 3808 AACAGGACTGCTTAGAAGGCACGAAGGCACACTACTGCTGGAATTTGTCTGA 3857  
 367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
 3858 CCTAGGCTACAGAGCTCTCTAAGAAGGCCAGATTTGCAGAGAGAGAGG 3907  
 384 alThrTyrLeuGlyTyrSerLeuArgGlyGlnArgTrpLeuThrGlu 400  
 3908 TAACATACCTTGGGTACAGTTTCGGGAGGGGAGGAGCATGGCTGAGGAG 3957  
 401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
 3958 GCACGGAAGAAACTCTAGTCCAGATACCGGCCCAACACACAGCCAAACA 4007  
 417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434  
 4008 AGTGAGAGAGTTTTGGGACAGCTGGAATTTTCAGACTGTGGATCCCGG 4057  
 434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
 4058 GGTTCGGACCTTAGCAGCCCTCTACCCACTAACCAAGAAAGAGG 4107  
 451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467  
 4108 GAATTCCTCTGGGCTCTCTGAGCACCAGAGGCAATTTGATGCTATCAAAA 4157

467 sAlaLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProP 484  
|||||  
4158 GCCCTGCTGAGCGACCTGCTCGGCCCTCCCTGCAGTAACTAAACCCCT 4207  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
|||||  
4208 TTACCCCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTTAAACC 4257  
501 GlnThrLeuGlyProThrArgProValAlaTyrLeuSerLysLysLe 517  
4258 CAAACTCTAGGACCATGGAGGAGACCTGTTGCCCTACCTGTCAAAGAAGCT 4307  
517 uAspProValAlaSerGlyThrProValCysLeuLysAlaIleAlaLav 534  
|||||  
4308 CGATCCCTGTAGCCAGTGGTTGGCCCGTATGCCCTCAAGGCTATCGCAGCTG 4357  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
|||||  
4358 TGCCATACCTGGTCAAGGACGCTGACAAATTGACTTTGGGACAGATATA 4407  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
|||||  
4408 ACTGTAATAGCCACCATCGTGTGGAGAACATCGTTCCGGCAGCCCCAGA 4457  
567 pArgTyrMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeu 584  
|||||  
4458 CCGATGATGACCAACCGCCGATGACCCACTATCAAAAGCCTGCTTCTCA 4507  
584 hrGluArgValThrPheAlaProAlaAlaLeuAsnProAlaThrLeu 600  
|||||  
4508 CAGAGAGGTCACGTTCCGCTCCACAGCCGCTCTCAACCTGCCACTCTT 4557  
601 LeuProGluThrThrAspGluProValThrHisAspCysHisGlnLeu 617  
4558 CTGCCTGAGAGACTGATGAACCAAGTCACTCATGATTGCCATCAACTATT 4607  
617 uIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4608 GATTGAGGAGACTGGGGTCCGGAAGGACCTTACAGACATACCGCTGACTG 4657  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
|||||  
4658 GAGAACTGTTAACTCTGTTCACTGACGGAAGCAGCTATGTGTGGAAAGT 4707  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
|||||  
4708 AAGAGATGGCTGGGGCGCGGTGTGTGACGGGACCCGACGATCTGGGC 4757  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4758 CAGCAGCCTGCCGGAAGGAACCTTACGACAAAGGCTGAGCTCATGGGCC 4807  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700  
|||||  
4808 TCACGCAAGCTTTGGCGCTGGCCGGAAGGAATCCATAAACATTTATACA 4857  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrI 717  
4858 GACAGCAGGTATGCCCTTTCCGACTGCACACGTACACGGGGCCATCTATAA 4907  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4908 GCAAAGGGGTGCTTACCCTCAGCAGGGGGAATAAAGAACAAAGAGG 4957  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4958 AAATCTAAGCCTATTAGAAGCCTTACATTTCCCAAAAAGGCTAGCTATT 5007  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
5008 ATACACTGTCTGGACATCAGAAAGCAAAAGATCCCATATCCAGAGGAA 5057  
767 nGlnMetAlaAspArgValAlaLysGlnAlaValAlaValAsnLeuL 784

|||||  
5058 CCAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGGGGTGTAACCTTC 5107  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
|||||  
5108 TGCTTATGATAGAAACACCAAGCCCAAGCCCGGACGACGAGTACAC 5157  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
|||||  
5158 CTAGAAGACTGGCAAGAGATAAAAGATAGACCAGTCTCTGAGACTCC 5207  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
5208 GGAAGGAGCTGCTATACCTCAGATGGGAAGCAAAATCTGCCCCACAAG 5257  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
5258 AAGGGTTAGAAATATGTCCAACAGATACATCGTCTAACCCACCTAGGA 5307  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
5308 AAACACCTGCACAGTGTGTCAGAACATCTCCTTATCATGTCTGAGGCT 5357  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 884  
5358 ACCAGAGTGGCTGATTGGTGGTCAAACTGTGTGCCCTGCCAGCTGG 5407  
884 alAsnAlaAsnProSerArgIleProGlyLysArgLeuArgGlySer 900  
5408 TTAATGCTAAATCTCTCCAGAAATACCTCCAGAAAGAGACTAAGGGGAA 5457  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
5458 CACCAGGCGCTCACTGGAAAGTGGACTTCACTGAGGTAAAGCCGGCTAA 5507  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
5508 ATACGGAACAAATATCTATTGGTTTTGTAGACACCTTTTCAGGATGGG 5557  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
5558 TAGAGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAGAAA 5607  
951 lleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlyse 967  
5608 ATACTGGAGGAAATTTTTCCAAAGATTTGGAAATACCTAAGGTAAAGGTC 5657  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
5658 AGACAATGGTCCAGCTTTCTGTTGCCAGGTAAAGTCAGGAGCTGGCCAAGA 5707  
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
5708 TATTTGGGATTTGTTGAAACTGCTATGTGCATACAGACCCCAAGGCTCA 5757  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
5758 GGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCCTTTACTAAAT 5807  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
5808 GACCGCGGAGACTGGCGTTAATGATTGATAGCTCTCCTGCCCTTTGTGC 5857  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
5858 TTTTTAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAA 5907  
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067  
5908 TTACTCTACGGGGACCCCCCATTTGGTGAATTTGCTTCCGTACATAG 5957  
1067 rAlaAspValLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084

5958 TGCTGACCTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGGCACCTTG 6007  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer... 1099  
|||||...  
6008 AGTGGGTGAGACACAGCGGTGGAGGCACTCCGGAGGCTACTCAGGA 6057  
1100 GlyGlyAspLeuGlnVal 1105  
|||||...  
6058 GGAGGAGACTTGCAGATC 6075

seq\_name: /SDS2/gcdata/geneseq/geneseq/NA1997.DAT.AAV09698

seq\_documentation\_block:

ID AAV09698 standard; DNA; 3320 BP.

AC AAV09698;

XX DT 19-MAY-1998 (first entry)

XX DE Porcine retrovirus pol and env DNA.

XX KW Porcine retrovirus; PoEV; POL protein; ENV protein; vaccine;  
diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX OS Porcine retrovirus.

XX FH Key Location/Qualifiers

FT CDS 23..2793

FT /\*tag= a

FT /product= POL protein

FT /note= "polymerase protein"

FT CDS 2642..3297

FT /\*tag= b

FT /product= ENV protein

FT /note= "envelope protein"

XX W09740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUTRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX Polynucleotide encoding porcine retrovirus expression product -

XX useful to develop products for use in vaccines, diagnosis and

XX xeno-transplantation

XX Claim 4; Fig 1; 69pp; English.

XX This DNA sequence encodes the porcine retrovirus (PoEV) polymerase (POL)  
XX and envelope (ENV) proteins. These proteins can be used to develop viral  
XX vaccines, antisense nucleic acids, ribozymes and other antiviral agents.  
XX They can also be used in xeno-transplantation technology and as  
XX diagnostic tools.

XX Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;

XX alignment\_scores:

XX Quality: 4840.00 Length: 921

XX Ratio: 5.261 Gaps: 0

XX Percent Similarity: 99.891 Percent Identity: 99.674

XX alignment\_block:

US-09-171-553B-5 x AAV09698

Align seg 1/1 to: AAV09698 from: 1 to: 3320

271 AspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAl 287  
|||||

22 GATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGC 71  
|||||

287 aPheGluTrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpT 304  
|||||

72 CTTTGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGA 121  
|||||

304 hrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla 320  
|||||

122 CCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAGGCC 171  
|||||

321 LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLe 337  
|||||

172 CTACACAGGACCTGCCAACTTCAGGATCCACACCCCTCAGGTGACCT 221  
|||||

337 uLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspC 354  
|||||

222 CTTCCAGTAGTGGATGACCTGCTTCTGCGGGAGCCACCAACAGGACT 271  
|||||

354 yLeuGluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGlyTyr 370  
|||||

272 GCTTAGAAGGTACGAAGGCACCTACTGCTGGAATGTCTGACCTAGGCTAC 321  
|||||

371 ArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLe 387  
|||||

322 AGAGCCTCTGCTAAGAAGCCCAAGATTTCAGAGAGAGAGTAACATACTT 371  
|||||

387 uGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysL 404  
|||||

372 GGGGTACAGTTTGGGGGGGGGCGGCGGATGGCTGACGGAGGCACGGAAGA 421  
|||||

404 yThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGlu 420  
|||||

422 AAACCTGTAGTCCAGATACCGGCCCAACACACAGCCAAACAAAGTGAGAG 471  
|||||

421 PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaTh 437  
|||||

472 TTTTGGGGACAGCTGGAATTTTCAGACACTGTGATCCCGGGGTTTGCAGC 521  
|||||

437 rLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyGlyPheSerT 454  
|||||

522 CTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAAAGGGGATCTCCT 571  
|||||

454 rAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuLeu 470  
|||||

572 GGGCTCTGAGCACACAGAGGCAATTTGATGCTATCAAAAGGCCCTGCTG 621  
|||||

471 SerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTy 487  
|||||

622 AGCGCACTCTCTGGCCCTCCCTGAGCTAACTAAACCCCTTTACCCCTTA 671  
|||||

487 rValAspGluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuG 504  
|||||

672 TGTGGATGAGCGTAAGGAGTAGCCCGAGGAGTTTAAACCCCAACCCCTAG 721  
|||||

504 lYProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProVal 520  
|||||

722 GACCATGGAGGAGACCTGTTCCTACCTGTCAAAGAGCTTGATCCTGTA 771  
|||||

521 AlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLe 537  
|||||

772 GCCAGTGGTTGGCCCGTATCTGAAGGCTATCGCAGCTGTGGCCACTACT 821  
|||||

537 uValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleA 554  
|||||

822 GGTCAAGGAGCGCTGACAAATTTGACTTTGGGACAGAAATATACCTGTAATAG 871  
|||||

554 laProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMet 570  
|||||

```

|||||
872 CCCCCCATGATTGGAGAACATCGTTCGGCAGCCCCCAGACCGATGGATG 921
571 ThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuThrGluArgVa 587
|||||
922 ACCAACCCCGCATGACCCACTATCAAAAGCGCTCTCTCACAGAGAGGGT 971
587 lThrPheAlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluG 604
|||||
972 CACTTTCGCTCCACCAGCCGCTCTCAACCGCTGCCACTCTCTCCCTGAAG 1021
604 luthrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluGlu 620
|||||
1022 AGACTGATGAACAGTGACTCATGATTCGCATCAACTATTGATTGAGGAG 1071
621 ThrGlyValArgLysAspLeuThrAspIleProLeuThrClyGluValLe 637
|||||
1072 ACTGGGTCGCCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCT 1121
637 uThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMeta 654
|||||
1122 AACCTGGTTCACTGACGGAAGCAGCTATGTGTTGGAAGTTAAGAGATGG 1171
654 laglyAlaAlaValValAspGlyThrArgThrIleThrPalaSerSerLeu 670
|||||
1172 CTGGGGCGCAGTGGTGGAGCGGACCGCACGATCTGGGCCAGCAGCTG 1221
671 ProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAl 687
|||||
1222 CCGGAGGAAGACTTCAGCGNAAGGCTGAGCTCATGGCCCTCACGCAAGC 1271
687 aLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgT 704
|||||
1272 TTTTCGGCTGGCCGAAGGAAATCCATAAACATTTATACGGACAGAGT 1321
704 yrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGly 720
|||||
1322 ATGCCCTTGGAGCTGCACAGCTACACGGGGCCATCTATAAACAAGGGGG 1371
721 LeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluLeuLeuSe 737
|||||
1372 TTCTTACCTCAGCAGGGAGGAAATAAGACAAAGAGAGAAATCTTAAG 1421
737 rLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysP 754
|||||
1422 CCTATTAGAAGCCTTACATTTGCCAAAAGGCTAGCTATTATACACTGTC 1471
754 roglyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMetAla 770
|||||
1472 CTGGCATCAGAAAGCCAAAGATCTCATATCTAGAGGGAACAGATGGCT 1521
771 AspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleI 787
|||||
1522 GACCGGGTTGCCAAGCAGCAGCCAGCTGTATTACCTTCTGCCCTATAAT 1571
787 eGluThrProLysAlaProGluProArgArgGlnTyrThrLeuGluAspT 804
|||||
1572 AGAAACGCCCAAGGCCCAAGACCCAGACGACAGTACACCTTAGAAGACT 1621
804 rpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThr 820
|||||
1622 GGCAGAGATAAAAAGATAGACCACTTCTCTGAGACTCCGGAGGGGACC 1671
821 CysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGl 837
|||||
1672 TGCATACCTCATATGGGAGGAATCTGCCCCCAAGAAAGAGGTTAGA 1721
837 uTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuG 854
|||||
1722 ATATGTCCAACAGATACATCGTCTAACCCACCTAGGAACCTAACACCTGC 1771
854 lncLnuValArgThrSerProTyrHisValLeuArgLeuProGlyVal 870
|||||

```

```

1772 AGCAGTTGGTCAGAACATCCCCCTTATCATGTTCTGAGGCTACCAGAGTG 1821
871 AlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAs 887
|||||
1822 GCTGACTCGTGTCAACATTTGTGTGCCCTGCCAGCTGGTTAATGCTAA 1871
887 nProSerArgIleProProGlyLysArgLeuArgGlySerHisProGlyA 904
|||||
1872 TCCTTCCAGAAATACCTCCAGGAAGAGACTAAGGGGAAGCCACCAGGCG 1921
904 laHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyAsn 920
|||||
1922 CTCACCTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAC 1971
921 LysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyr 937
|||||
1972 AAATATCTATTGGTGTGTTGTAGACACCTTTTCAGGATGGTAGAGCTTA 2021
937 rProThrLysLysGluThrSerThrValValAlaLysLysLysIleLeuGluG 954
|||||
2022 TCCCTACTAAGAAAGAGACTTCAACCGTGTGGCTAAGAAATACTTGGAGG 2071
954 luIlePheProArgPheGlyIleProLysValIleGlySerAspAsnGly 970
|||||
2072 AAATTTTCCAAAGATTTGGAATACCTAAGGTAATAGGCTCAGACAATGGT 2121
971 ProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIle 987
|||||
2122 CCAGCTTTCGTTGCCAGGTAAAGTCAGGAGCTGGCCAAAGATATTGGGGAT 2171
987 eaSrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValG 1004
|||||
2172 TCATTGGAAACTGCATTTGTGCATACAGACCCCAAGCTCAGACAGGTAG 2221
1004 luArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrGlu 1020
|||||
2222 AGAGGATGAATAGAACCATTAAGAGAGACCTTACCATAATTGACCACAGAG 2271
1021 ThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgVa 1037
|||||
2272 ACTGCATTAATGATTTGATGGCTCTCTGCCCTTTGTGCTTTTAGGGT 2321
1037 lArgAsnThrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrG 1054
|||||
2322 GAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAATTGCTCTACG 2371
1054 lyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAspVal 1070
|||||
2372 GGGGACCCCGCTTGGCAGAAATTTGCCATATAGTGTCTGATGTG 2421
1071 LeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValAr 1087
|||||
2422 CTGCTTTCCAGCCCTTGTCTTAGGCTCAAGGGCTCGAGTGGGTGAG 2471
1087 gGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGlyGlyAspLeuG 1104
|||||
2472 GCAGCGAGCGTGAAGCAGCTCCGGAGGCCCTACTCAGGAGGAGACTTGC 2521
1104 lncValProHisArgPheGlnValGlyAspSerValTyrValArgHis 1120
|||||
2522 AAGTTCACATCGCTTCCAGTTGGAGATTTCAGTCTATGTTAGACGCCAC 2571
1121 ArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLe 1137
|||||
2572 CGTCAGGAAACCTCGGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTT 2621
1137 uThrThrProThrAlaValLysValGluGlyIleProThrTrpIleHisA 1154
|||||
2622 GACCACACCAACGCTGTGAAGTGCAGAGGAATCCCCACCTGGATCCATG 2671
1154 laSerHisValLysProAlaProProProAspSerGlyTrpArgAlaGlu 1170
|||||
2672 CATCCACGTTAAGCCGCGCCACCTCCCGATTTCGGGGTGGAAAGCGGAA 2721

```

1171 LysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSe 1187  
|||||  
2722 AAGACTGAAATCCCTTACCTTCGCTCCATCGGTGCTTACTC 2771  
1187 rAsnAsnSer 1191  
|||||  
2772 TGTCATAACTCC 2784  
seq\_name: /SDS2/gcdata/geneseq/geneseq/NA2001.DAT.AAF77726  
seq\_documentation\_block:  
ID AAF77726 standard; DNA; 7333 BP.  
XX  
AC AAF77726;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Defective retroviral genome isolated from PK-15 cell line.  
XX  
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;  
KW ss.  
XX  
OS Unidentified.  
XX  
PN US6190861-B1.  
XX  
PD 20-FEB-2001.  
XX  
PF 13-DEC-1996; 9605-0766528.  
XX  
PR 14-DEC-1995; 9505-0572645.  
XX  
PA (GEO ) GEN HOSPITAL CORP.  
XX  
PI Fishman JA;  
XX  
DR WPI; 2001-256211/26.  
DR P-PSDB; AAB73282, AAB73283, AAB73284.  
XX  
PT Assessing risk of endogenous retroviruses in clinical practice and in  
PT xenotransplantation, comprises using probe sequences derived from swine  
PT or miniature swine retroviral genome -  
XX  
PS Claim 1; Fig 2; 127pp; English.  
XX  
CC The present invention relates to a method for screening a cell or tissue  
CC for the presence or expression of a retrovirus (RV), comprising  
CC contacting a target nucleic acid from the cell or tissue with a second  
CC nucleic acid from the present invention (e.g. the present sequence or a  
CC fragment thereof). The method is useful for RV detection and to assess  
CC graft transplantation risk. Screening of animals allows the elimination  
CC of donors with active replication of known viruses. Inactive proviruses  
CC can be detected and inactivated, allowing identification and elimination  
CC of potential human pathogens derived from swine in a manner not possible  
CC in the outbred human organ donor population and is important to the  
CC development of human xenotransplantation.  
XX  
SQ Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

2173 GGAGAGCGGGTTCGAGCCCCCTCCCGAGCCAGGGTAACCTTTGAAGGT 2222  
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34  
|||||  
2223 GGAGGGCAACACAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTCAG 2272  
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50  
|||||  
2273 TGCTGCTACAACCATTTAGGAAACTAAAAAGAAAAAATCCTGGGTGATG 2322  
51 GlyAlaThrGlyGlnArgGlnTyrProTyrThrArgArgThrValAs 67  
|||||  
2323 GGTGCCACAGGGCAACGGCAGATATCCATGGACTACCCGAGAACCGTTGA 2372  
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84  
|||||  
2373 CTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTTCATCCTCGTGC 2422  
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100  
|||||  
2423 CAGTACCCCTTCTAGGTAGACTTACTGACCAAGATGGGAGCTCAAT 2472  
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleTh 117  
|||||  
2473 TCITTTGACCAAGAGACAGAGAGTGTCTGTGAATAACAACCCATCAC 2522  
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134  
|||||  
2523 TGTGTTGACCTCCAATTTAGATGATGAATATCGACTATATTTCTCCCAAG 2572  
134 allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150  
|||||  
2573 TAAAGCCTGATCAAGATATACAGTCTGTTGGAGCAGTTTCCCAAGCC 2622  
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
|||||  
2623 TGGGCAGAAACCCGAGGGATGGGTTGGCAACGAAAGTCCCCACAGGT 2672  
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
|||||  
2673 TATTCACCTGAAGCCAGTGTCTACACAGTATCAGTCAGACAGTACCCCT 2722  
184 euSerArgGluAlaArgGluGlyIleTyrProHisValGlnArgLeuIle 200  
|||||  
2723 TGAGTAGAGAGGCTCAGAGGAATTTGGCCGATGTTCAAGATTAAATC 2772  
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217  
|||||  
2773 CAACAGGGCATCCTAGTTCCTGTCCAATCCCTTGGAAATACCTCCCTGCT 2822  
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
|||||  
2823 ACCGGTTAGGAAGCCTGGGACCAATGATATCCAGCAGTACAGGACTTGA 2872  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
|||||  
2873 GAGAGGTCAATAAAGGGTGCAGAGACATACACCCAGCGTCCCGAACCT 2922  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValle 267  
|||||  
2923 TATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAACCTGGTACACAGTATT 2972  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
|||||  
2973 GGACTTAAAGATGCCCTTCTTCTCCCTGAGATTTACACCCCACTAGCAAC 3022  
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300  
|||||  
3023 CACTTTTACCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAG 3072  
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
|||||  
3073 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGCGACCATCTT 3122

alignment\_scores:  
Quality: 4489.50 Length: 1191  
Ratio: 5.033 Gaps: 2  
Percent Similarity: 74.895 Percent Identity: 74.643  
alignment\_block:  
US-09-171-553B-5 x AAF77726 ..  
Align seg 1/1 to: AAF77726 from: 1 to: 7333  
1 GlyArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17  
|||||



```

317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
|||||
3123 TGAGGAAGCCCTACACAGGACCTGGCCAACTTCAGGATCCAAACCCCTC 3172
334 lnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
|||||
3173 AGGTGACCCCTCCAGTACGTGATGACCTGCTTCTGGGGAGCCACC 3222
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAs 367
|||||
3223 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTGCTGA 3272
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
|||||
3273 CCTAGGCTACAGAGCCCTCTCTAAGAAGGCCACAGATTTCAGAGAGAGG 3322
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
|||||
3323 TAACATACCTTGGGGTACAGTTTCGGGGGGGGCAGCGATGGCTGACGGAG 3372
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGI 417
|||||
3373 GCACGGAAGAAACTGTAGTCCAGATACCGCCCCCAACACACGCCAAACA 3422
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
|||||
3423 AGTGAGAGAGCTTTTGGGACACCTGGATTTTCACACCTGTGGATCCCGG 3472
434 lyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
|||||
3473 GGTTCGACCTTAGCAGCCCCACTCTACCGGCTAACCAAGAAAAAGGG 3522
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
|||
3523 GGTG..... 3527
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3527 ..... 3527
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3527 ..... 3527
501 GlnThrLeuGlyProThrPargArgProValAlaTyrLeuSerLysLysLe 517
3527 ..... 3527
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlav 534
3527 ..... 3527
534 aAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550
3527 ..... 3527
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567
3527 ..... 3527
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584
3527 ..... 3527
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
3527 ..... 3527
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617
3527 ..... 3527
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634
3527 ..... 3527
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650
3527 ..... 3527
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667
3527 ..... 3527
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAla 684
3527 ..... 3527
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
3527 ..... 3527
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717
3527 ..... 3527
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734
|||||
3528 .....CTTACTCTCAGCAGGGAGGAAATAAGAACAAAGAGG 3564
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750
|||||
3565 AAATCTTAAGCCTATTAGAAGCCTTACATTTGCCAAAAAGGCTAGCTATT 3614
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767
|||||
3615 ATACACTGCTCTGGACATCAGAAAAAGCCAAAGATCTCATATCTAGAGGAA 3664
767 nGlnMetAlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuL 784
|||||
3665 CCAGATGGCTGACCGGTTGCCAAGCAGCAGCCCGAGGCTGTAACTTC 3714
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800
|||||
3715 TGCTCTAATAAGAAAGCCCAAGCCCAAGCCAGACAGACAGTACACC 3764
801 LeuGluAspTrpGlnGluIleLysLysLysLeuAspGlnPheSerGluThrPr 817
|||||
3765 CTAGAAGACTGGCAGAGATAAAAAAGATAGACCATCTCTGAGAGTCC 3814
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834
|||||
3815 GGAGGGGACCTGTATACCTCATATGGGAAGAAATCTGCCCCCAAG 3864
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850
|||||
3865 AAGGTTAGAATATGTCCAACAGATACATCGCTTAACCCACCTAGGAAC 3914
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867
|||||
3915 AAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTCTGAGGCT 3964
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeuV 884
|||||
3965 ACCAGGAGTGGCTGACTCGGTGGTCAACATTTGTGCCCTGCCAGCTGG 4014
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
|||||
4015 TTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAGC 4064
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917
|||||
4065 CACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCGCGCTAA 4114
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934

```

```

|||||
4115 ATACGGAACAATAATCTATTGGTTTGTAGACACCTTTTCAGGATGG 4164
934 aGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950
4165 TAGAGGCTTATCCCTACATAAAAGAGACTTCAACCGTGGTGGCTAAGAAA 4214
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967
4215 ATACTGGAGGAATTTTCCAGATTGGAATACCTAAGGTAATAGGGTC 4264
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984
4265 AGACAATGGTCCAGCTTTCGTGCCAGTAAGTACAGGAGCTGCCCAAGA 4314
984 leLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
4315 TATTGGGGATTGATTGAAACTGTCATTTGTCATACAGACCCCAAGCTCA 4364
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017
4365 GCACAGGTAGAGAGATGAATAGAACCATTAAGAGAGACCTTACCAATT 4414
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034
4415 CACCACAGAGACTGGCATTAATGATTGGATGGCTCTCCTGCCCTTTGTGC 4464
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050
4465 TTTTAGGGTGAGGAACACCCCTGGACAGTTTGGGCGTACCCCTATATAA 4514
1051 LeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSe 1067
4515 TTGCTCTACGGGGACCCCGCTTGGCAGAAATTCCTTTGCACATAG 4564
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084
4565 TGCTGATGTGCTTTTCCAGCCTTTGTTCTTAGGCTCAAGCGCTCG 4614
1084 luTpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100
4615 AGTGGTGAGGACGAGCGTGAAGCAGCTCCGGAGGCGCTACTCAGGA 4664
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117
4665 GGAGACTTGCAAGTTCCACATCGCTTCCAGTTGGAGATTTCAGTCTATGT 4714
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134
4715 TAGACGCCACCGTCGAGGAACCTCGAGACTCGGTAGAAGGACCTTATC 4764
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150
4765 TCGTACTTTTGACACACCAACGCGTGTGAAGTCGAAGG..... 4804
1151 TrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTr 1167
4804 ..... 4804
1167 pArgAlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuV 1184
4805 .....AATPCCCTTAAGCTTCGCCCTCATCGCGTGG 4835
1184 aProTyrSerAsnAsnSer 1191
4836 TTCCTTACTCTGTCAATAACTCC 4858

```

seq\_name: /SIDS2/gcdata/geneseq/geneseqn/NA1997.DAT.AAT74883

seq\_documentation\_block:

ID AAT74883 standard; cDNA: 7393 BP.

XX

AC AAT74883;

```

XX 09-FEB-1998 (first entry)
DT Porcine retrovirus cDNA (defective).
XX
DE Retrovirus; porcine; GAG protein; POL protein; ENV protein;
XX xenotransplantation; Infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
KW Porcine retrovirus.
XX
XX Key Location/Qualifiers
FT CDS 598..2172
FT /*tag= a
FT mat_peptide 598..2169
FT /*tag= b
FT /*note= "putative GAG protein"
FT CDS 2320..4737
FT /*tag= c
FT /*note= "putative POL coding region (partial) as
FT described in the specification"
FT mat_peptide 2320..3522
FT /*tag= d
FT /*note= "putative POL protein (partial)"
FT mat_peptide 3516..4326
FT /*tag= e
FT /*note= "putative POL protein (partial)"
FT mat_peptide 4332..4748
FT /*tag= f
FT CDS 4738..6725
FT /*tag= g
FT /*note= "putative ENV coding region (partial) as
FT described in the specification"
FT mat_peptide 4752..6722
FT /*tag= h
FT /*note= "ENV protein (partial)"
XX
XX WO9721836-A1.
XX
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX
XX 14-DEC-1995; 95US-0572645.
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Fishman JA;
XX
XX WPI: 1997-332804/30.
XX P-PSDB; AAW32091-W32095.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX viruses in transplant or other tissue and for assessing risk of
XX transmitting infection to graft recipient
XX
XX Claim 16; Fig 2; 128pp; English.
XX
XX This cDNA sequence represents a defective purified swine retrovirus
XX found in PK-15 cells containing the putative coding regions for viral
XX GAG, POL and ENV proteins. There are a few in frame stop codons and
XX apparent frame shifts in the given coding sequence which alter features
XX of the translation. This sequence and PCR fragments generated from the
XX presence of porcine retroviruses prior to xenotransplantation.
XX Transplantation can increase the likelihood of retroviral activation if
XX intact and infectious proviruses are present. The porcine retroviral
XX sequence can be used to generate probes to determine the level (e.g.
XX copy number) of intact (i.e. potentially replicating) porcine provirus
XX sequences in a strain of xenograft transplantation donors. It can be
XX used to detect mutations, genetic lesions or viral recombinants and to
XX determine the histological localisation of activated retrovirus. Using
XX Polymerase Chain Reaction DNA Quantitation (PQO) on blood mononuclear

```

CC cells, infectivity titration and susceptibility testing can be  
 CC performed. Ultimately animal donors without intact porcine retroviral  
 CC sequences or a lower copy number of viral elements could be selected.  
 XX  
 SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

## alignment\_scores:

Quality: 4489.50 Length: 1191  
 Ratio: 5.033 Gaps: 2  
 Percent Similarity: 74.895 Percent Identity: 74.643

## alignment\_block:

US-09-171-553B-5 x AAT74883 ..

Align seg 1/1 to: AAT74883 from: 1 to: 7393

```

1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysVa 17
2173 GGGAGACGGGGTTCGGACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGT 2222
17 lGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerV 34
2223 GAGGGGCAACAGTTGAGTTCTCTGGTTGATACCGGAGCGGAGCATTCAG 2272
34 alLeuLeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMet 50
2273 TGCTGTACAAACATTAGGAAACTAAAGAAAAAATCTCTGGGTGATG 2322
51 GlyAlaThrGlyGlnArgGlnTyrProTrpThrArgArgThrValAs 67
2323 GGTGCCACAGGCAAGCGAGTAGTATCCATGGACTACCCGAAGAACCGTTGA 2372
67 pLeuGlyValGlyArgValThrHisSerPheLeuValIleProGluCysP 84
2373 CTGGGAGTGGGACGGGTAAACCACCTGCTTCTGGTCATCCCTGAGTGCC 2422
84 roValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
2423 CAGTACCCCTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAAAT 2472
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIle 117
2473 TCTTTTGAACAGGAAGAACAGCAAGTGTCTGTAATAACAACCCATCAC 2522
117 rValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnV 134
2523 TGTGTTGACCCCTCCAATTAGATGATGAATATCGACTATATTCCTCCCAAG 2572
134 allysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAla 150
2573 TAAAGCCTGATCAAGATATACAGTCTGTTGGAGCAGTTTCCCAAGCC 2622
151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167
2623 TGGGCAGAAACCGCAGGGATGGGTTTGGCAAGCAAGTTCCCCACAGGT 2672
167 lIleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184
2673 TATTCAACTGAAGCCAGGTGCTACACAGTATTCAGTCAGACAGTACCCCT 2722
184 euSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
2723 TGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAGATTAAATC 2772
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLe 217
2773 CAACAGGGCATCTAGTCTCTGTCCTCAATCCCTTGAATACTCCCTGCT 2822
217 uProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234
2823 ACCGGTTAGGAAGCCCTGGGACCAATGATTATCGACCAGGACTTGA 2872

```

```

234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250
2873 GAGAGTCAATAAAGGGTGCAGGACATACACCAACGGTCCCGAACCTT 2922
251 TyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValIle 267
2923 TATAACCTCTTGAGCGCCCTCCCGCTGAACGGAACCTGGTACACAGTATT 2972
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284
2973 GGACTTAAAGATGCTTCTCTGCTGAGATTACACCCCTAGCCCAAC 3022
284 roLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
3023 CACTTTTACCTTCGAATGGAGATCCAGGTACGGGAAGAACCGGGCAG 3072
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317
3073 CTCACCTGGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTT 3122
317 eAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334
3123 TCACCAAGCCCTACACAGGACCTGCCCAACTTCAGGATCCAACACCTC 3172
334 InValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThr 350
3173 AGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 3222
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuSerAs 367
3223 AAACAGGACTGCTTAGAAGGTACGAAGGCACCTACTGCTGGAATTGCTGA 3272
367 pLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384
3273 CTTAGCTACAGAGCCTCTGCTAAGAGGCCAGATTTCGAGGAGAGAGG 3322
384 alThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
3323 TAAACATACTGGGGTACAGTTTGGCGGCGGCGAGCATGGCTGACGAGG 3372
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGl 417
3373 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACAGCCAAACA 3422
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProG 434
3423 AGTGAGAGAGTTTTTGGGACAGCTGGATTTTGACAGCTGGATCCCGG 3472
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450
3473 GGTTCGAGACCTTAGCAGCCCACTACCGCTAACCAAGAAAGAAAGGG 3522
451 GlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467
3523 GGTTCG..... 3527
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484
3527 ..... 3527
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
3527 ..... 3527
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLe 517
3527 ..... 3527
517 uAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaLav 534
3527 ..... 3527
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550

```

3527 ..... 3527  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
3527 ..... 3527  
567 pArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
3527 ..... 3527  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
3527 ..... 3527  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
3527 ..... 3527  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
3527 ..... 3527  
634 lyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGluGly 650  
3527 ..... 3527  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAl 667  
3527 ..... 3527  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
3527 ..... 3527  
684 euThrGlnAlaLeuArgLeuAlaGluLysSerIleAsnIleTyrThr 700  
3527 ..... 3527  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLy 717  
3527 ..... 3527  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
3528 .....CTTACCTCAGCAGGAGGAGAAATAAGAACAAAGAGG 3564  
734 luIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
3565 AAATTCAGCCATTAGAACCCCTTACATTGGCCAAAAGGCTAGCTATT 3614  
751 IleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAs 767  
3615 ATACACTGTCTGGACATCAGAAAGCCAAAGACTCATATCTAGAGGGAA 3664  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
3665 CCACATGGCTGACCGGGTTGCCAAGCAGGCAGCCAGCGCTGTTAAACCTTC 3714  
784 euProIleIleGluThrProLysAlaProGluProArgArgGlnTyrThr 800  
3715 TGCCTATAATAAGAACCCCAAGCCCAAGAACCCAGACAGACAGTACACC 3764  
801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrPr 817  
3765 CTAGAACACTGGCAGAGATATAAAGATAGACACAGTCTCTCAGACTCC 3814  
817 oGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysG 834  
3815 GGAGGGACCTGCTATACCTCATATGGGAAGGAATCCTGCCCCACAAAG 3864  
834 luGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThr 850  
|||||

3865 AAGGGTTAGATATGTCCAACAGATACATCGTCTAAACCACCTAGGAAC 3914  
851 LysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLe 867  
|||||  
3915 AAACACCTGCAGCAGTGTGTCAGAACATCCCTTATCATGTCTCAGGCT 3964  
867 uProGlyValAlaAspSerValValLysHisCysValProCysGlnLeu 884  
|||||  
3965 ACCAGAGTGGCTGACTCGGTGGTCAAAACATTGTGTGCCCTGCCAGCTGG 4014  
884 alAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900  
|||||  
4015 TTAATGCTTAATCCTTCCAGATACCTCCAGAAAGAGACTAAGGGGAAGC 4064  
901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLy 917  
|||||  
4065 CACCCAGGCGCTCACTGGGAAGTGCATCTCACTGAGGTAAAGCCGGCTAA 4114  
917 sTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpV 934  
|||||  
4115 ATACGGAACAAATATCTATTGGTGTGTAGACACCTTTTCAGGATGGG 4164  
934 alGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysLys 950  
|||||  
4165 TAGAGCTTATCCTACTAAAAAGAGACTTCAACCGTGTGTGCTAAGAA 4214  
951 IleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySe 967  
|||||  
4215 ATACTGGAGGAAATTTTCCAAGATTGGAAATACCTAAGGTAATAGGCTC 4264  
967 rAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysI 984  
|||||  
4265 AGACAAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGGACTGGCCAAGA 4314  
984 leuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000  
|||||  
4315 TATTGGGATTGATTGAAAACATGTCATTGTGCATACAGACCCCAAGCTCA 4364  
1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLe 1017  
|||||  
4365 GCACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACCATAAT 4414  
1017 uThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValL 1034  
|||||  
4415 GACCACAGACTGGCATTAATGATTGGATGGCTCTCTCGCCCTTTGTGC 4464  
1034 euPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlu 1050  
|||||  
4465 TTTTGGGTGAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATATAA 4514  
1051 LeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSe 1067  
|||||  
4515 TTGCTCTAGGGGGACCCCCCGTTGGCAGAAATTCCTTTGCACATAG 4564  
1067 rAlaAspValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuG 1084  
|||||  
4565 TGCTGATGTGCTGCTTCCAGCCCTTGTCTCTAGGCTCAAGGCGCTCG 4614  
1084 luTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerGly 1100  
|||||  
4615 AGTGGGTAGGCAGGAGCGTGGAGAGCTCCGGGAGCCCTACTCAGGA 4664  
1101 GlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrVa 1117  
|||||  
4665 GGAGACTTGCAGATTCCACATCGCTTCCAAAGTTGCAGATTTCAGTCTATGT 4714  
1117 lArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrL 1134  
|||||  
4715 TAGAGCCACCGCTGCAGGAACCTCGAGACTCGTAGAGGACCTTATC 4764  
1134 euValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThr 1150  
|||||  
4765 TCGTACTTTTGACCAACACACCGGCTGTGAAGTCGAAGG..... 4804



151 TrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnVa 167  
2738 TGGCGGAACAAGCAGGATGGGTAAACAAGTGCCTCCAGTCGT 2787  
167 IleGlnLeuLysAlaSerAlaThrProValSerValArgGlnTyrProL 184  
1788 GGTAGAAGTAAAGCTGATGATCCCATCCATCAGTAAAGCAATACCCCA 2837  
184 euSerArgGluAlaArgGluGlyIleTrrProHisValGlnArgLeuIle 200  
2838 TGACGAAGAAGCTAAGAGGGCATCCGACCTCATATTCGGAGGCTGCTA 2887  
201 GlnGlnGlyIleLeuValProValGlnSerProTrrPsnThrProLeuLe 217  
2888 GACCAAGGAATTTAGTGGCTGTGAGTCCCTGGAACATCCACATTTT 2937  
217 uproValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuA 234  
2938 GCCAGTAGAAGCAGGAGCAATGATTATCGCCCGGTCAAGACCTCC 2987  
234 rgGluValAsnLysArgValGlnAspIleHisProThrValProAsnPro 250  
2988 GGAAGTAAATAAAGGCTCTGCACATTCACCTACAGTCCCGAACCCG 3037  
251 TyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrrTyrThrValLe 267  
3038 TACAATTTACTAAGCTCTCCACCCGAGAGAACCTGGTATACGGTCCT 3087  
267 uAspLeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnP 284  
3088 GGATTTAAAGATGCTTCTTTTGGCTGCGTCTGCACCCCAAGAGTCAAC 3137  
284 roLeuPheAlaPheGluTrrArgAspProGlyThrGlyArgThrGlyGln 300  
3138 TTCGTGTTGGCTTGAATGGAGAACCCAGAGGGCGGACAGACTGGTCAA 3187  
301 LeuThrTrrPThrArgLeuProGlnGlyPheLysAsnSerProThrIlePh 317  
3188 TTAACCTGGACTAGACTACCAAGGGGTTCAAAACTCCCTACCTGTT 3237  
317 easpGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProG 334  
3238 TGACGAAGCCCTCCATCGGGATTAGCACCTTTCGGCGCTCAAAACCCCC 3287  
334 InValThrLeuLeuGlnTrrValAspAspLeuLeuAlaGlyAlaThr 350  
3288 AGCTTACCCTACTGCAGTACGTAGATGATCTTTAATCGCAGCGCCTCA 3337  
351 LysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuSerAs 367  
3338 AAAGAGCTATGTCAACAGGGGACTGAGAGACTCCTCAAGAACTGGGGAA 3387  
367 pleuGlyTrrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluV 384  
3388 TTGGGGGTATCAGTTTCGGCTAAAGGCACAAATTTGTCAAACTGAGG 3437  
384 alThrTyrLeuGlyTrrSerLeuArgGlyGlyGlnArgTrrPleuThrGlu 400  
3438 TAATCTACTGGGGTATCTTTGCGAGGAGGATAAAGATGCTCCTCACTGAA 3487  
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysG 417  
3488 GCCCGGAAAAAGACTGTCTATGATGATTCACCGCCCAACACACACCGCA 3537  
417 nValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrrPleProG 434  
3538 GGTACGTGAGTTCTGGGGACTGCTGGCTTTGTAGACTCTGGATTCCAG 3587  
434 lypPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGly 450  
3588 GCTTTGCGACCTAGCAGCACCCCTATATCCCTTACTAGGGAAGGAATT 3637  
451 GlyPheSerTrrAlaProGluHisGlnLysAlaPheAspAlaIleLysLy 467

3638 CCTTTGAATGGAAGAAACACCAAGAGCTTTTGAGGCTATCAAAATC 3687  
467 sAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspValThrLysProp 484  
3688 GTCTCTAATGACTGCCCTGCGCTAGCATTTACAGACTTGACTAAGTCT 3737  
484 heThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500  
3738 TCGTCTCTATATGTGGACGAGAGCGGCATAGCCAGAGGGGTGCTGACA 3787  
501 GlnThrLeuGlyProTrrPargArgProValAlaTyrLeuSerLysLysLe 517  
3788 CAAGCACTAGGACCTTGAAGAGACCGTAGCTATCTGTCAAAAAAAT 3837  
517 uAspProValAlaLaserGlyTrrProValCysLeuLysAlaIleAlaAlaV 534  
3838 GGATCCCGTGTAGTGGATGGCCACATGTCTGAAGGCTATTGCAGCAG 3887  
534 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIle 550  
3888 TAGCCCTGTGATCAAAAGATGCTGATAAACTGACAAATGGGACAGAGGTG 3937  
551 ThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAs 567  
3938 ACTGTCGTGGCCCTCATGCCCTGGAGATTTGTGCGGACGCCACTGA 3987  
567 pAspTrrMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuT 584  
3988 TAGATGGATGACTAATGCCCGGATGACACACTATCAGAGTCTGCTGATA 4037  
584 hrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600  
4038 ATGACCGGGTAACTTTGCCCCCTGCCATTTCTCAACCCAGCACCCCTC 4087  
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLe 617  
4088 CTCCTCTTAAGAAATGATTCCTGCCAGTACATCGATGTGCAGACATCT 4137  
617 uIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrG 634  
4138 GGCTGAAGAAATTTGGGACCAAGAACCTGACTGACCAACCTGGCCTG 4187  
634 lyGluValLeuThrTrrPheThrAspGlySerSerTyrValValGluGly 650  
4188 GA...GCTCTAGTGGTACACGATGGCAGCAGTCTTCTGATTGAGGA 4234  
651 LysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrrAl 667  
4235 AAGCGAGGGCTGGAGCTGGGTGGTGGATGGAAAAAGGTAAATTGGGC 4284  
667 aSerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaL 684  
4285 AAGTGCCTTGCCTGAAGAACTTCGCGACAGAAAGGCTGAACATAGCGC 4334  
684 euThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTrrThr 700  
4335 TTACAGAGCCCTCCGAGAGGCGCAAGGTAAGATCAATTAATTTACACT 4384  
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTrrLy 717  
4385 GACAGCGGTATGCTTTGCTACTGCACATATCCATGGGGCCATCTACAG 4434  
717 sGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluG 734  
4435 GCAGCGAGGGTATTGACTTCAGCCGTAAGACATCAAAACAAAGAAAG 4484  
734 lulleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIle 750  
4485 AAATTTGCGCCCTTAGAGCCATACATCGCCCAAGAGGTAGCCATC 4534  
751 ileHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLys 767

4535 ATCCACTGTCCTCCGCTCACCAGAAAGGAGAGGAGTCTAGTGGCCCAAGGCCAA 4584  
767 nGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuL 784  
4585 CCGGATGGCAGACTCAGTACCAACACAGGTGTCCAGGGGCCATGATCT 4634  
784 euProIleIleGluThrPro...LysAlaProGluProArgArgGlnTyr 799  
4635 TAACGTAAAGAAATCCCTCCAAAAGCCCTGAG..... 4669  
800 ThrLeuCluAspPrpGlnGlnIleLysLysIleAspGlnPheSerGluTh 816  
4670 .....GATGAAACTATGATATATAAAGAACTATTTGGAC 4704  
816 rProGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisL 833  
4705 TAGTGATCCCTCCCATCTCTTTTCGAAGGAAAAATAGACTTCACACCCG 4754  
833 ysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 849  
4755 AGGAAGGAATAAATTTGTCAAAGGACTACACCACTTTACCCACCTGGGA 4804  
850 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuAr 866  
4805 GTTGAAAAATGATGAGACTGATTAATAAGTCCCGGTATCAAGTCCCTAA 4854  
866 gLeuProGlyValAlaAspSerValValLysHisCysValProCysGlnL 883  
4855 CTGGAAGTCTAGTGGCTCAAAAGATTAATAACTCTCGAAAGCGTGTGCAT 4904  
883 euValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGly 899  
4905 TCACATATGCAACTAAAACTACAAAGACCTGGAAAGAGACACACGGGA 4954  
900 SerHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAl 916  
4955 GACCGTCTCGAGTGTATGGGAGGTGGATTTTACTGAAGTTAAACCCGG 5004  
916 aLysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyT 933  
5005 AATGTATGGTAATAAGTATCTGTAGTATTTGTAGACACCTTTTCAGGAT 5054  
933 rpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 949  
5055 GGGTTGAAGCGTCCCCACTAAAACTGAGACTGCCAGATTGTGGCCAAG 5104  
950 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGl 966  
5105 AAGATTTTGAAGAAATCCCTGCCAAGATATGGAGTACCTAAGGTAATCGG 5154  
966 ySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaL 983  
5155 GTCCGACATGGACACCGCTTTGTGCCAGGTAAAGTCAGGATTTGCCCA 5204  
983 ysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSer 999  
5205 CTCAGTTGGGCATTGATTGGAATTTACATTGTCTTACCCCTCAAGC 5254  
1000 SerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLy 1016  
5255 TCAGGACAGGTAGAGGATGAATAGAACATTAAGAGGACCTTTGACTAA 5304  
1016 sLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheV 1033  
5305 ATTTGGCCATCGAGACCGGGGAAAGACTGGGTGGCTCTCCTCCCTCTTG 5354  
1033 alLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1049  
5355 CGCTCTCCGCGCCGGAACACCCCGGAGGTTTGGGCTCAGCTCCTTTT 5404  
1050 GluLeuLeuTyrGlyGlyProProProLeuAlaGlu.....IleAl 1063  
5405 GAAAGTTCTGTATGGGGACCTCCCTCTTTAATAAAGATGGTGGAAACATT 5454

1063 aPheAlaHisSerAlaAspValLeuSerGlnProLeuPheSerArgL 1080  
5455 GGTTCCTCCGATTCAGGCTCTGCTTACCCTCTCT...TTGCTTATTCATT 5501  
1080 euLysAlaLeuGluTyrValArgGlnArgAlaTyrLysGlnLeuArgL 1096  
5502 TAAAGCCCTGAAAGTAAATTAGGACCCAAATTTGGGACAGCTAAAGACG 5551  
1097 AlaTyrSerGlyGlyAspLeuGlnValProHisArgPheGlnValGlyAs 1113  
5552 GCCTACACCCCGAGGACACCCAGCTACCCACGAGTTCAGGTCGGGA 5601  
1113 pSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArgTyrPL 1130  
5602 CCAAGTCTTGTCAGACGACATCGAACCGGTAGCCTTGAACACACCGTGA 5651  
1130 ysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGlu 1146  
5652 AGGGACCTTATTTAGTGTACTTAACAACCTCTTACGGCAGTGAAGTTGAC 5701  
1147 GlyIleProThrTrpIleHisAlaSerHisValLysProAlaProProPr 1163  
5702 GGGATTGCTCTCTGGATCCACGCTTCACACGTCAGAGGGCCCCAGTCA 5751  
1163 oAsp.....SerGlyTrpArgAlaGluLysThrGluAsn 1175  
5752 AGATGAAGAAACCCAGGAGACAACTGGGCTGTGAGCCACTGATAACC 5801  
1175 rLeuLysLeuArgLeuHisArgLeuValProTyrSerAsnAsnSer 1191  
5802 CTCTTAAGCTTCGCTGCGTCGAGGAGGCCCTC.....CATCAC 5842  
1192 ProGly 1193  
5843 CCTGGA 5848









```

284 snThrProArgAsnSerProGluValProValIlyThrGluIleuArgLeu 300
|||||
6470 ACACGGCTTAACAACTCCCGAGGTGCTGCTTAAAGACAGAGAGAACTC 6519
301 PheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspProAs 317
|||||
6520 TTCAGTCTCATCCAGGAGGCTTCCAAAGCCATCACTCACAGACCTTGA 6569
317 PAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrTyrG 334
|||||
6570 TGGCCTCTCTCTGTTGGCTTGTCTATCCCTCAGGCGCTCTTATTATG 6619
334 IuGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsn 350
|||||
6620 AGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGAGCATTAAGAT 6669
351 GlnCysThrTrpGlySerArgAsnLysLeuThrLeuGluValSerG 367
|||||
6670 CAATGTACATGGGGGTCCGAAATTAAGCTTACCCCTCAAGTTCGCG 6719
367 yLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCys 384
|||||
6720 GAAGGGGACATGATAGAGAAACCTCCCATCCACCAACACCTTTGCT 6769
384 yTrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuValPro 400
|||||
6770 ATAGTACTGTGCTTATAGACAGGCTCAGAAATCAGTATTAGTACCT 6819
401 GLyTyrAsnArGTrPrAlaCysAsnThrGlyLeuThrProCysValSe 417
|||||
6820 GGTATTAACAGTGGTGGCATCAATACGTGGTTAACCCCTGTGTTTC 6869
417 rThSerValPheAsnGlnSerLysAspPheCysPheValMetValGlnIle 434
|||||
6870 CACTCAGTCTCAACCAATCAAGATTTCTGTGTCATGTGTCCTCAATG 6919
434 alProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAsp 450
|||||
6920 TCCCGGAGTGTACTACCATCTGAGSAGATGGTCTGTGATGAATATGAC 6969
451 TyrArGTrAsnArGProLysArGluProValSerLeuThrLeuAlaVal 467
|||||
6970 TATCGGTATTAACGACCAAAAGAGAACCCGTATCCCTTACCTAGCTGT 7019
467 lMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaVal 484
|||||
7020 AATGCTGGATTAGGAGCGGCGCTGGCGTAGAACAAGGAGACGCTGCC 7069
484 euIleThrGlyProGlnGlnLeuGlnLysGlyLeuGlyLysLeuHisAla 500
|||||
7070 TGATCAGACAGACACAGCAGCTAGAGAAAGACTTGGTGTAGCTAATGCG 7119
501 AlMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLeuG 517
|||||
7120 GCCATGACAGAAATCTCCAGGCTTAAAGAGAGCTGTGAGCAACCTAGA 7169
517 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArGArgLyl 534
|||||
7170 AGAGTCCCTGACTCTTTGCTGTGAGTGTCTTACACAGACGAGGGGAT 7219
534 euAspLeuLeuPheLeuArgGluGlyLysLeuCysAlaAlaLeuLysGlu 550
|||||
7220 TAGATCTGCTGTCTTCTAAGAGAAAGTGGTATGTGACAGCTTAAAGAA 7269
551 GluLysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAs 567
|||||
7270 GAAATGCTCTTATGTAGATCACTCAGAGCCATTCAGAGACTCCATGAA 7319
567 nLysLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584
|||||
7320 CAAGCTTAGAAAGAGTTAGAGAGCGCTCGAAGGAGAAAGAGGCTGACC 7369
584 lngIlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLeu 600

```

```

|||||
7370 AGGGGTGTTGAAGATGTTCAACAGAGTCTCCTTGATGACACACCTG 7419
601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThrVa 617
|||||
7420 CTTTGTCTGTGACGGGCGCCCTAGTACTGCTGCTGTTACTTACAGT 7469
617 lGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgVal 634
|||||
7470 TGGGCTTCTCTTAATTATAGTGTGTGCTTGTGTAGAGACAGTGA 7519
634 erAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSer 650
|||||
7520 GTCCAGTCAGATCATGTACTTACGCAACAGTACCAAGGCTTCTGAGC 7569
651 GlnGlyGluThrAspLeu 656
|||||
7570 CAAGGAGAAACATGACCTC 7587

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV09699
seq_documentation_block:
ID AAV09699 standard; DNA: 8196 BP.
XX
AC AAV09699;
XX
DT 19-MAY-1998 (first entry)
XX
DE Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS Porcine retrovirus.
XX
FH Key
FH Location/Qualifiers
FT CDS 576..2126
FT /tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
FT CDS 2143..5733
FT /tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
FT specification"
FT CDS 5606..7576
FT /tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX
PN W09740167-A1.
XX
PD 30-OCT-1997.
XX
PE 18-APR-1997; 97WO-GB01087.
XX
PR 10-FEB-1997; 97GB-0002668.
PR 19-APR-1996; 96GB-0008164.
XX
PA (IMUT-) IMUTRAN LTD.
PA (OONE-) O-ONE BIOTECH LTD.
XX
PI Galbraith DN, Haworth C, Lees CM, Smith KT;
XX
DR WPI: 1997-535851/49.
XX
PT Polynucleotide encoding porcine retrovirus expression product -
PT useful to develop products for use in vaccines, diagnosis and
PT xeno-transplantation
XX
PS Claim 4; Fig 2; 69pp; English.
XX
CC This DNA sequence encodes the porcine retrovirus (POEV) virion core

```

CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These  
CC proteins can be used to develop viral vaccines, antisense nucleic acids,  
CC ribozymes and other antiviral agents. They can also be used in  
CC xeno-transplantation technology and as diagnostic tools.

SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

alignment\_scores:

Quality:	3509.00	Length:	656
Ratio:	5.357	Gaps:	0
Percent Similarity:	99.848	Percent Identity:	99.848

```
alignment_block:
```

US-09-171-553B-6 x AAV09699

Align seg 1/1 to: AAV09699 from: 1 to: 8196

1	MeHisProThrIleuSerArgGHisLeuProThrArgIylgiLupr	17
5606	ATGCATCCACACGTTAAAGCYGGCCACACTCCCGACTCGGGGTGGAGAGCC	5655
17	oLysArgLeuArgIleProIleuSerPheAlaSerIleAlaIrrPheLeuT	34
5656	GAAAGACACGAGAAATCCCCCTTAAAGCTTCGGCTCCATGCTCGTGGTCTTA	5705
34	hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer	50
5706	CTCTAACATATACTCCCGCAGGCGACGTAAGTAACGCCCTATATAGACAGCTCG	5755
51	AsnProHisArgProIleuSerIleuThrPleuIleIleAspProAspTh	67
5756	AAACCCCATAGACTTATTCCTTACCTCGCGATTAATGACCTCGATAC	5805
67	rGlyValThrValAsnSerThrArgGlyValAlaProArgIylThArgTrp	84
5806	GGGTGTCACTGTAATATAGACATCGAGGGTGTGTCTCTAAGAGGACCTGGT	5855
84	rProGlnLeuHisPheCysLeuAlaGleuIleAsnProAlaValLysSer	100
5856	GGCGCTGAAGTGCATTTGCCCCCATGTGATTAACCCCCCTGTTAARAGC	5905
101	ThrProAsnLeuValArgSerIyrgIyPheIyrcyCysProGlyTh	117
5906	ACACTCCCAACACTAGTCCGTAAGTATGAGGTCTATATGCTGCCACGAC	5955
117	rGluLysGluLysIyrcyGlyGlySerGlyLysIleuSerPheCysArgArgT	134
5956	AGAAAGAGAGAAATACTGTGGGGGTTCTGGGGAATCCTCTGAGAGAT	6005
134	rPserCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu	150
6006	GGACTCGCTACCTCCACACATGGAGACTGGAAATGGCGCACTCTCTTC	6055
151	GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTrpIyry	167
6056	CAGGACCGGGTAAATTCCTCTTGTCAATTCGGCCCCGGGGAAGTACAA	6105
167	SMetMetLysLeuIyryLysAspLysSerCysSerProSerAspLeuAspTr	184
6106	AATGATGAAACTATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT	6155
184	ylLeuLysIleSerPheThrGluArgLysTrnGlyLysTySerLysVal	200
6156	ATCTAAGATATAGTTCTACGAAAGGAAACAGGAAATATTCAAAAGG	6205
201	AspLysTrpIyrgIuLeuGlyAsnSerPheLeuLeuIyrgIylgiLylAl	217
6206	GATTAATGATATAGCTGGGGAATAGTTTTTATATATATGCGGGGAGAC	6255
217	agLysSerThrLeuThrIleArgLeuArgIleGluThrGlyIyrgLuprop	234
6256	AGGTCACACTTAAACATTTGCCCTTAGATAGAACGGGACCAAAACCC	6305

234	royaIaIameGlyProAspIysValIleuAlaGluGlnGlyProProAla	250
6306	CTTGtGGCAATGGGAGCCGGATAAATAGCTGGCTGAACAAGGGGCCCGCCGCC	6355
251	LeuAluProProHIsAsnIleuProValProGlnIleuThrSerIeuArgPr	267
6356	CTGGAGGCCACCGCATTAACCTTCCGGGTGCCCAATTAACTCGTGGCGGCC	6405
267	oAspIleThrGlnProProSerAsnSerThrThyGlyLeuIleProThra	284
6406	TGACATTAACACAGCCGCGCTGACACAGACAGACACTGATGTATTCCTACCA	6455
284	snThrProArgAsnSerProGlyValProValIysThrGlnArgLeu	300
6456	ACACGGCTTAAGAACTCCCGAGGGTTCCTGTTAAGACAGGACAGAGACTC	6505
301	PhSerIeuIleGlnGlnValAlaPheGlnAlaIleAsnSerThrAspProAs	317
6506	TTTACGTCTCATCCAGGAGCTTTTCCAAGCATCAACTCCACCCAGCTCGA	6555
317	paIaThrSerSerCysThrPheCysIeuSerSerGlyProProThyArg	334
6556	TGCCACTTCTCTTGTGTGGCTTGTGTATCCCTCAGGGCCCTCTATTATG	6605
334	IuGlyMetAlaIysGluAlaGlyPheAsnValThrIysGluHisArgAsn	350
6606	AGGGGATGGCTAAAGAAAGAAATTCATATGTGACCAAGAGGCTATGAAT	6655
351	GlnCysThrThrProIysSerArgAsnIysIleuThrIleuThrGlnValSerG	367
6656	CAATGTACATGGGGGCTCCGGAATTAAGCTTACCCTCACTGAATTTTCGG	6705
367	ylYsgIyThrCysIleGlyLysAlaProProSerHisGlnHisIleuCysT	384
6706	GAAAGGAGACATCATATAGAAAGAAAGCTCCCATGCCACCAACACTTTGCT	6755
384	ylSerThrValValylGlyGluGlnAlaSerGluAsnGlnIyIleuValPro	400
6756	TAATGACTGTGGTTTATATGAGAGGCTCTCAAAATCAGATTATGATACCTT	6805
401	GlyTyrrAsnArgTrpTrpAlaCysAsnThrGlyIleuThrProCysAlaSe	417
6806	GGTATTACAGGTGGTGGCATATCAATCTGGGTTAAACCCCTGGTCTTC	6855
417	rThrSerValPheAsnGlnSerIysAspPheCysValIleValGlnIleY	434
6856	CACCTCAGCTTCAACCAATCCAAAGATTCTGTGTCATGATGCCAATATGG	6905
434	alProArgValIyTyrrHisProGluGluValValIleuAspGluTyrrAsp	450
6906	TCCCCCGAGTGTATACCATCTCGTAGGAAGTGGTCTTGATGAATATGAC	6955
451	TyrrArgTyrrAsnArgProIysArgIleuProValSerIeuThrIleuAlaVa	467
6956	TATGGGTATTAACCGACCAAAAGAAAGAACCCGATATCCCTTACCTGAGCTG	7005
467	IleIleuGlnIleuGlyThrAlaValAlaGlyValGlyThrGlyThrAlaAla	484
7006	AATGCTCGAGATTAGGAGAGCGCGTGGGCTAGGAACAGGACAGCTGCC	7055
484	eulIeThrGlyProGlnIleuGlnIulysGlyIleuGlyIleuHisAla	500
7056	TGATCCACAGGACCAAGAGCTTAAGAAAGAGACTTGGAGACTCATATCGG	7105
501	AlaMetThrGluAspIleuArgAlaIleuysIleuSerValSerAsnIleuG	517
7106	GCCATTGACAGAAATCTCCAGGCTTTAAAGAGATCTGTTACCAACTAGA	7155
517	uGluSerIeuThrSerIeuSerGluValIleuGlnAsnArgArgGlyL	534
7156	AGAGTCCCTCACTTCTGTCTGAAGTGTTCTTACAGAACCGGAGGGGAT	7205

534 euAspLeuLeuPheLeuArgGluGlyLeuCysAlaLeuLysGlu 550  
 |||||||  
 7206 TAGATCTGCTTCTTAAGAGAGGTGGTATATGTCAGCCTTAAGAA 7255  
 551 GluCysCysPheTyrValAspHisSerGlyAlaLeuArgSerMetAs 567  
 |||||||  
 7256 GAATGTGCTTCTATGATGATCAGCAGAGCCATCAAGACTCATAA 7305  
 567 nlyLeuArgLysLysLeuGluArgArgArgGluArgGluAlaAspG 584  
 |||||||  
 7306 CAAGCTTGAAGAAAGTTAGAGAGCCGTGAGAGGAAAGAGGCTGACC 7355  
 584 lnglyTyrPheGluGlyTyrPheAsnArgSerProTyrMetThrLeu 600  
 |||||||  
 7356 AAGGGGTGGTTGAGAGATGGTTCAACAGGTCTCCCTTGATGACCACTG 7405  
 601 LeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThy 617  
 |||||||  
 7406 CTTTCTGCTGTGACGGGCCCCCTAGTACTGCTGCTGTTACTACAGT 7455  
 617 lglTyrCysLeuLeuAsnArgPheValAlaPheValArgGluArgValS 634  
 |||||||  
 7456 TGGGCTTGTCTTAATTAATAGTTGTGCTTGTAGAGAGCAGTGA 7505  
 634 eAlaValGlnIleMetValLeuArgGlnIleTyrGlnGlyLeuLeuSer 650  
 |||||||  
 7506 GTGCAGTCCAGATCATGTACTTAGCAACAGTACCAAGGCTTCTGAGC 7555  
 651 gInglyGluThrAspLeu 656  
 |||||||  
 7556 CAAGAGAAACTGACCTC 7573

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.AAV09703

seq\_documentation\_block:

ID AAV09703 standard; DNA: 1974 BP.

XX AAV09703;

DT 20-MAY-1998 (first entry)

DE Porcine retrovirus Raji clone ENV DNA.

KW Porcine retrovirus; POEV; ENV protein; envelope protein; vaccine;

KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT CDS 1..1974

FT /\*tag= a

FT /product= ENV protein

FT /note= "envelope protein"

XX MO9740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97MO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUPRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI: 1997-535851/49.

XX P-PSDB: AAW39274.

XX Polynucleotide encoding porcine retrovirus expression product -

PT useful to develop products for use in vaccines, diagnosis and

PT xeno-transplantation  
 XX  
 PS Claim 6; Fig 4; 69pp; English.  
 XX  
 CC This sequence encodes the porcine retrovirus envelope (ENV) protein  
 CC isolated from the human cell line Raji. Such viral proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.  
 XX  
 SQ Sequence 1974 BP; 534 A; 489 C; 483 G; 468 T; 0 other;

alignment\_scores:  
 Quality: 3455.00 Length: 658  
 Ratio: 5.291 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 99.088

alignment\_block:  
 US-09-171-553B-6 x AAV09703 ..

Align seg 1/1 to: AAV09703 from: 1 to: 1974

1 MethisProthrLeuSerArgArgHisLeuProThrArgGlyGluPrt 17  
 |||||||  
 1 ATGCATCCACAGTTAACTGACCTGGCCACCTCCGACTCGGGGTGAGAGCC 50  
 17 olyArgLeuAlaGlyLeuProLeuSerPheAlaSerIleAlaTyrPheLeu 34  
 |||||||  
 51 GAAAGACCTGAGAAATCCCTTAAAGCTTGGCTCCATCGCTGATGCTCTTA 100  
 34 hrLeuthrIleThrProGlnAlaSerSerLysArgLeuLeuAspSerSer 50  
 |||||||  
 101 CTCTAACATTAACCTCCAGGCGCAGTGTAAAGCCTTATAGACAGCTCG 150  
 51 AsnProHisArgProLeuSerLeuThrTrpLeuIleLeuAspProAsp 67  
 |||||||  
 151 AACCCCATAGACCTTATCCCTGACCTGCTGATTAAGACCTGATAC 200  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgIleThrTpt 84  
 |||||||  
 201 GGGTGTCACTGTAAATAGACCTGAGGTGTGCTCCATAGAGGACCTGGT 250  
 84 rPrGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 |||||||  
 251 GGCTGAACATGCAATTTCTGCTCCGATTTGATTAACCCCGCTTAAAGC 300  
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 |||||||  
 301 ACACCTCCCAACCTAGTCCGATGTTAGGTGCTTATGCTGACCCAGGAC 350  
 117 rGlnLeuGlnLysTyrCysGlyGlySerGlyGlySerPheCysArgArgT 134  
 |||||||  
 351 AGAGAAAGCAAAATATCTGGGGGTCTGGGGAATCCCTCTGTAGAGAT 400  
 134 rPserCysValThrSerAsnAspGlyAspTyrLysTyrProLysLeu 150  
 |||||||  
 401 GGAGCTGGTACCTCCACAGATGAGACCTGGAATGCGCATCTCTCTC 450  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
 |||||||  
 451 CAGGACCGGGTAAATATCTCTTGTCAATTCGCGCCGCGCAAGTCAAA 500  
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAsp 184  
 |||||||  
 501 AATGATGAACATATATTAAGATAGAGCTGCTCCCATCAGACTTACATT 550  
 184 YrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLysV 200  
 |||||||  
 551 ATCTAAAGATTAAGTTTCACTGCAAAAGAAAGCAAGGAAATATTCAAAAG 600  
 200 alaPlyStrPyrGluLeuGlyAsnSerPheLeu..LeuTyrGlyGlyG 216  
 |||||||

601 TGCATTAATGCTATGACCTGGCAATAGTTTTTATATATGCGGGG 650  
 216 YALAGYSerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 233  
 651 AGAGGGTCCACTTAAACATTCGCTTAGATAGAGGGGAGCAAGAC 700  
 233 ROProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 249  
 701 CCGCTGGGCAATGGGACCGGATTAAGTACTGGCTGACAGGAGGCGCCCG 750  
 250 AlAlaGluProProHisAsnLeuProValProGlnLeuThrSerLeuArg 266  
 751 GCCCTGGAGCCACCGCATTAAGTTCGCGGTCGCCCAATTAACCTCGCTGG 800  
 266 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 283  
 801 GCCTGACATTAACACAGCGGCTTAGCAACAGTACCAGTGGATTATTCCTA 850  
 283 hAsnThrProArgAsnSerProGlyValProValLysThrGlyLysArg 299  
 851 CCAACAGCGCTAGAAACTCCCGAGGTTCCTGTAAAGACAGAGACAGAGA 900  
 300 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 316  
 901 CTCTTCAAGTCTCATCCAGGAGCTTTCAGGCCATCACTCCACGACGCC 950  
 316 oAspAlaThrSerSerCysTrpLeuGlySerLeuSerSerGlyProProTyr 333  
 951 TGATGCCACTTCTTCTTGGCTTGTGTATCCTCAGGGCCCTTATT 1000  
 333 YrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 349  
 1001 ATGAGGGAGGAGGCTAAAGAAAGAAATCAATGTGACCAAGAGCATAGA 1050  
 350 AsnGluCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValG 366  
 1051 AATCAATGATCATGGGGGTCCGAAATTAAGCTTACCCTCAGAAATTTTC 1100  
 366 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383  
 1101 CGGGAAAGGGACATGCATAGGAAGAAAGCTCCCATCCACCAACACTTT 1150  
 383 YsTrpSerThrValValYrGluGlnAlaSerGluAsnGlnTrpLeuVal 399  
 1151 GCTATGTACTGTGTTATGACAGCGCTCAGAAATATCAGTATTAGTA 1200  
 400 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysTa 416  
 1201 CCTGGTTATACAGGTGGTGGCATCAATACTGGGTAAACCCCTGTGT 1250  
 416 lSerThrSerValPheAsnGlnSerLysAspPheCysValMetValGlnI 433  
 1251 TTCACCTGAGTCTTCACCAATCCAAAGATTGTGTGTCATGTCGTCANA 1300  
 433 lValAlaProArgValIYrYrHisProGluGluValValLeuAspGluTr 449  
 1301 TCTCTCCCGAGTGTACTACCATCTGAGGAAGTGTCTCTGATGAAATAT 1350  
 450 AspTyrArgTrpAsnArgProLysArgGluProValSerLeuThrLeuAl 466  
 1351 GACTATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCTTAGC 1400  
 466 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483  
 1401 TGTAATGCTCGATTTAGGGAGCGGCTTGGCGTAGAGAACAGGACAGCTG 1450  
 483 lAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyLeuHis 499  
 1451 CCCTGATACAGAGCACACAGACCTAGAGAAAGAGACTGTGATGATACAT 1500  
 500 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerSole 516  
 1501 GCGGCCATGACAGAGATCTCCAGGCTTAAAGAGAGTGTGTAACAACCT 1550

516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533  
 1551 AGAAGAGTCCCTGACTCTTGTGTGAAGTGTTCTTACAGAACCGAGGG 1600  
 533 lYLeuAspLeuLeuPheLeuArgGlyGlyGlyLeuCysAlaAlaLeuLys 549  
 1601 GATTAGATCTGCTGTTCTTAAGAGAGGTGGGTATATGACGCCCTTAAA 1650  
 550 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSer 566  
 1651 GAGAATGTGCTTCTTATGTAGATACCTCAGAGCCATCAGAGACTCAT 1700  
 566 lAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583  
 1701 GAACAGAGCTTGAAGAAAAGTTAGAGAGCGTCGAAAGGAAAGAGGCTG 1750  
 583 sPGLnglyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 599  
 1751 ACCAGGGGTGGTTTGAAGATGTTCAACAGGTCTCCTTGGATGACACCC 1800  
 600 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuTh 616  
 1801 CTGCTTCTGCTCTGACGGGCGCCCTAGTACTGCTCTCTGTTACTTAC 1850  
 616 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgLysArg 633  
 1851 AGTTGGGCTTGCTTAATTAAGTGTGTTGCTTGTAGAGACGAG 1900  
 633 aSerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeuLeu 649  
 1901 TGAGTGCAGTCCAGATCATGTGACTTAGCAACAGTACCAAGGCGCTCTG 1950  
 650 SerGlnGlyGluThrAspLeu 656  
 1951 AGCCAAGAGAAACTGACCTC 1971

seq\_name: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV82749  
 seq\_documentation\_block:  
 ID AAV82749 standard; DNA: 3482 BP.  
 AC AAV82749;  
 XX  
 DT 25-FEB-1999 (first entry)  
 XX  
 DE Pig endogenous retrovirus (PERV) - B envelope (env) gene region.  
 XX  
 KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
 KW probe; primer; detection; retrovirus; human tissue; xenotransplant;  
 KW primary porcine tissue; human cell line; porcine cell line; ss.  
 XX  
 OS Pig endogenous retrovirus.  
 XX  
 FH Location/Qualifiers  
 FT CDS 911..2884  
 FT /\*tag= a  
 FT /product= envelope\_protein  
 XX  
 PN W09853104-A2.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PE 18-MAY-1998; 98MO-GB01428.  
 XX  
 PR 16-MAY-1997; 97GB-0010154.  
 XX  
 PA (MEDIC-) MEDICAL RES COUNCIL.  
 XX  
 PI Sloye JP, Weiss RA;  
 XX  
 DR WPI; 1999-045324/04.  
 DR P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridizing to either  
 PT the PERV-A or PERV-B env gene - useful in the detection of  
 PT retroviruses, and their subtypes, in a sample of porcine/human  
 PT tissue

PS Claim 3: Page 21-23; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-B  
 CC envelope protein. PERV exists in two different subtypes, PERV-A and  
 CC PERV-B. The differences are reflected in sequence divergence in the  
 CC envelope genes. Probes and primers can be derived from the envelope  
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
 CC a method to detect retroviruses in a sample of porcine/human tissue,  
 CC particularly primary porcine tissue and human cell lines that have been  
 CC cultured in the presence of a porcine cell line, or human tissue from  
 CC one of the PERV env genes can also be determined.

SQ Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

# alignment\_scores:

Quality: 3433.00 Length: 658  
 Ratio: 5.257 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 98.176

## alignment\_block:

US-09-171-553b-6 x AAV82749 ..

Align seg 1/1 to: AAV82749 from: 1 to: 3482

1 MethisProthrleuSerArgrhIleuproThrArgGlyGluPr 17  
 911 ATGCATGCCAGCTTAAGCTGGGCCACCTCCGAGATCGGGTGGAGACC 960  
 17 oluSArgleuarGlieProleuSerPheAlaSerIlealatrPheleut 34  
 961 GAAAGAGCTGAGAAATCCCTTAAGCTTCGCCATCGCTGGTTCCTTA 1010  
 34 hleuThrIlethrProGlnAlaSerIleuArgleuIleasPserSer 50  
 1011 CTCTACATATACCTCCAGGCCAGTAGTAACGCTTATAGACAGCTCG 1060  
 51 AsnProHISArgProleuSerIleuThrPleuIleIleasProAspTh 67  
 1061 AACCCCAATAGACCTTTATCCCTTACCTGCGATGATTATGACCTGATAC 1110  
 67 rGlyValThrValAlaSerThrArgrGlyValAlaProArgrGlyThrTrpT 84  
 1111 GGGTGTCACTGTAATATGACACTCGAGGTGTGCTCTAGAGGACACTGT 1160  
 84 rPProGlnleuHISPheCysleuArgleuIleasProAlaValIlySer 100  
 1161 GGCCTGGAACGATTTGCTGCTCGATTGATTAAACCCGCTGTAAAGC 1210  
 101 ThrProAsnleuValArgSerThrArgGlyPheThrCysCysProGlyTh 117  
 1211 ACACCTCCCAACCTAGTCCGATGATTAGGTTCTATTGCTCCACAGCACC 1260  
 117 rGlyGlyGluIlySerThrCysGlySerGlyGluSerPheCysArgrTrp 134  
 1261 AGAGAAAGAGAAATACGTGTGGGGTCTGGGGAATCCTTCTGTGAGGAT 1310  
 134 rPserCysValThrSerAsnAspGlyAspTrpIlySTPProIleSerleu 150  
 1311 GGAAGTGGTCACTCCACAGATGAGAGCTGGAATGCGCATCTCTCTC 1360  
 151 GlnAspArgValIlySPheSerPheValAlaSerGlyProGlyIlySTYrly 167  
 1361 CAGGACCGGGTAAATTCCTTTGTCAATTCGCGCCCGGACAGTACAA 1410  
 167 SmeMetIlyleuIlyrlyAspIlySerCysSerProSerAspIleuAspTr 184

1411 AGTATGTAATATATATTAAGATTAAGAGCTGCTCCCATACAGACTTAGATT 1460  
 184 YrleuIlyIleSerPheThrGlu..ArgIySThrGlyIlySTySerIlySV 200  
 1461 ATCTAAAGATTAAGTTTCACTGAAAAGAGAAACAGAAAATATTCAAAAG 1510  
 200 AlaSplySTrPThrGlnleuGlyAsnSerPheIleu..LeuYrGlyGly 216  
 1511 TGATAAATGTATATGAGCTGGGGAATAGTTTATTAATATGGCGGGG 1560  
 216 YAlaGlySerThrleuThrIleArGleuArgIleGluThrGlyThrGluPr 233  
 1561 AGCAGGCTCCACTTAAACCATTCCTTAGATAGACAGGGGACAGAC 1610  
 233 roProValAlaMetGlyProAspIlyValleuAlaGluGlnGlyProPro 249  
 1611 CCCTGTGGGAGTGGGACCCGATAAAGTACTGCTGAAACAGGGGCCCCCG 1660  
 250 AlaIleuGluPrProProHISAsnleuProValProGlnleuThrSerleuAr 266  
 1661 GCCCTGAGACCCGATACCTTCGCGTCCCAATTAACCTCGTGGC 1710  
 266 qProAspIleThrGlnProProSerAsnSerThrThrGlyleuIleProT 283  
 1711 GCCTGACATATACACAGCCGCTACCAACGATACACACTGATGATTCCTTA 1760  
 283 hrAsnThrProArgAsnSerProGlyValProValIlySThrGlyGlnArg 299  
 1761 CCAACAGCGCTAGAACTCCCGAGGTGTCTCTGTTAAGACAGACAGAGA 1810  
 300 leuPheSerleuIleGlnGlyAlaPheGlnAlaIleasSerThrAspTr 316  
 1811 CTCTTCAGTCTCATCGAGAGCTTTCAGACCATCAATCCACCGCAC 1860  
 316 oAspAlaThrSerSerCysTrpIleuSerSerSerGlyProProTrpTr 333  
 1861 TGATGCCACTTCTCTGTGGCTGTGTCTATCTCATCGGGCCCTCTATT 1910  
 333 YrGlyGlyMetAlaIlyGluArgIlySPheAsnValThrIlyGlnHISArg 349  
 1911 ATGAGGGGATGGCTAAAGAGAAATTCATATGACCAAAAGACATAGA 1960  
 350 AsnGlnCysThrTrpGlySerArgrAsnIlyleuThrleuThrGlnIlyAla 366  
 1961 AATCAATGTCATGAGGGGCTCCGAAATTAAGCTTACCTCACTGAAGTTTC 2010  
 366 rGlyIlyGlyThrCysIleGlyIlyAlaProProSerHISGlnHISleu 383  
 2011 CGGGAAGGGACATGCATAGAAAGCTCCCATCCACACACACTT 2060  
 383 ySTySerThrValIlyrGluGlnAlaSerGlnAsnGlnIlyrleuVal 399  
 2061 GCTATATGACTGTGCTTATGACAGCCCTCAAGAAATCAGTATTATAGTA 2110  
 400 ProGlyTyraAsnArgrTrpTrpAlaCysAspThrGlyleuThrProCysVa 416  
 2111 CTTGTTATACAGGTGGTGGCATGCAATACGAGGTAAACCCCTGTGT 2160  
 416 IserThrSerValPheAsnGlnSerIlyAspPheCysValMetValGlnI 433  
 2161 TTCACACTGAGTCTTACACCAATCCAAAGATTTCTGTGCATATGTCMAA 2210  
 433 IeValProArgrValIlyrYrHISProGluGlnValIleuAspGlnIlyr 449  
 2211 TCGTCCCGCCAGTGTACTACATCTGAGGAATGTCTTGTATTAATAT 2260  
 450 AspTrArgrTyraAsnArgrProIlyAsnArgIlyProValSerleuThrleuAl 466  
 2261 GACTATCGGTATACCGACCAAAAGAAAGAACCCGTATCCCTTACCTTAC 2310  
 466 aValMetleuGlyleuGlyThrAlaValAlaGlyThrGlyThrAlaA 483

```

2311 TGTATGCTCGGATTAGGAGCGCGCTTGCGCTAGAGAAACAGGACGACTG 2360
483 1aleuIIeThrglyProglInleuGluYsgLYleuGLyLleuHIs 499
|||||
2361 CCTGATACAGAGACACGACGACTAGAGAAAGACTTGGTGACTTCAT 2410
500 A1a1aMeThrgluAspLeuArgAlaLeuYsgLusSerValSerAsnLe 516
|||||
2411 GCGGCGATGACAGAGATCTCCGAGCCTTAGAGAGATCTGTTTACCACT 2460
516 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533
|||||
2461 AGAAGAGTCCCTGACTCTTGTCTGAGTGTCTGAGTGTCTGACAGAACCGAGGG 2510
533 1LeuAspLeuLeuPheLeuAArgGluGlyLeuCysAla1a1LeuYs 549
|||||
2511 GATTAGATCTGCTGTTTCTTAGAGAGGTGGTTATGTGCGCCTTAA 2560
550 GluGluCysCysPheTyrValAspHisSerGlyAla1a1earGAspSerMe 566
|||||
2561 GAAGATGTGCTCTTATGTAGATCAGCTCAGAGACCATCAGAGACTCCAT 2610
566 tAsnYsLeuAArgYsLYsLeuGluAArgAArgAArgGluAArgGluAla 583
|||||
2611 GAGCAAGCTTAgAGAAAGGTTAGAGAGCGCTCGAAGGAAAGAGAGCGCTG 2660
583 spgInGlyTTPheGluGlyTTPPheAsnAArgSerProTTPMetThr 599
|||||
2661 ACCAGAGGTGGTTGAAGATGTTCAACAGCTCTCCTTGATGACAC 2710
600 LeuLeuSerAlaLeuThrglyProLeuValValLeuLeuLeuLeuTh 616
|||||
2711 CTGCTTCTGCTGACGAGCCCTAGTACTCTGCTGCTGCTTACTTAC 2760
616 rValGlyProCysLeuLeuAsnAArgPheValAlaPheValAArgGluAArg 633
|||||
2761 AGTTGGCGCTTGCTTAATATAGTTTGTGCTTGTGTTTGAAGAACGAG 2810
633 a1SerAlaValGlnLeuMetValLeuAArgGlnGlnTyrGlnGlyLeuLeu 649
|||||
2811 TGAGTGACAGTCCAGATCATGTACTTAGGCAACAGTACCAAGCGCTTCTG 2860
650 SerGlnGlyGluThrAspLeu 656
|||||
2861 AGCCAAGAGAAACTGACCTC 2881
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.AAV82748
seq_documentation_block:
ID AAV82748 standard; DNA; 2462 BP.
XX
AC AAV82748;
XX
DT 25-FEB-1999 (first entry)
XX
DE pig endogenous retrovirus (PERV)-A envelope (env) gene region.
XX
KM pig endogenous retrovirus: PERV-A; envelope protein: PERV-B; subtype:
KW probe; detection: retrovirus; human tissue; xenotransplant;
KM primary porcine tissue; human cell line; porcine cell line; ss.
XX
OS pig endogenous retrovirus.
XX
FH key location/Qualifiers
FT CDS 211..2193
FT /*tag= a
FT /product= envelope_protein
XX
PN W09853104-A2.
XX
PD 26-NOV-1998.
XX
PF 18-MAY-1998; 98WO-GB01428.

```

```

XX 16-MAY-1997; 97GB-0010154.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX Stoye JP, Weiss RA;
XX
XX WPI, 1999-045324/04.
XX
XX P-PSDB: AAW85453.
XX
PT Newly isolated nucleic acid probe capable of hybridizing to either
PT the PERV-A or PERV-B env gene - useful in the detection of
PT retroviruses, and their subtypes, in a sample of porcine/human
PT tissue
XX
XX Claim 6; Page 20-21; 36pp; English.
XX
XX The present sequence encodes a pig endogenous retrovirus (PERV)-A
XX envelope protein. PERV exists in two different subtypes, PERV-A and
XX PERV-B. The differences are reflected in sequence divergence in the
XX envelope genes. Probes and primers can be derived from the envelope
XX (env) genes of PERV-A and PERV-B. The probes and primers are used in
XX a method to detect retroviruses in a sample of porcine/human tissue,
XX particularly primary porcine tissue and human cell lines that have been
XX cultivated in the presence of a porcine cell line, or human tissue from
XX a patient with a xenotransplant. Subtype of PERV in a sample containing
XX one of the PERV env genes can also be determined.
XX
XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;
XX

```

```

alignment_scores:
Quality: 2594.50 Length: 680
Ratio: 4.420 Gaps: 11
Percent Similarity: 86.324 Percent Identity: 73.088

```

```

alignment_block:
US-09-171-553B-6 x AAV82748 ..

```

```

Align seg 1/1 to: AAV82748 from: 1 to: 2462

```

```

1 MethisProthLeuSerAArgHisLeuProthAArgGlyGlyLeu 17
|||||
211 ATGCATCCACGCTTAAGCCGCGCCACCTCCGATTCGGGGTGAAAGCC 260
17 oLysAArgLeuAArgIleProLeuSerPheAlaSerIleAlaTTPheLeu 34
|||||
261 GAAAGACTGAAATCCCTTAGCTTCCGCTCCATCGCGTGGTCTTAA 310
34 hIeuthrIleThrProGlnAlaSerSerLysAArgLeuLeuAspSerSer 50
|||||
311 CTCTGCAATTAATCTCCTCAAGTTAATGTTAAAGCTTGTGACACCCG 360
51 AspProHisAArgProLeuSerLeuThrTTPLeuIle1aAspProAspH 67
|||||
361 AATCCCATTAACCTTATCTCTCACCCTGTTACTTACGACCCGCTAC 410
67 rGlyValThValAsnSerThAArgIyValAlaProAArgIyThTPT 84
|||||
411 AGGTATTAATATTAACAGCACTCAAGGAGGCTCCCTTGAGCACTGCT 460
84 rPProGluLeuHisPheCysLeuAArgLeuIleAsnProAlaVal..... 98
|||||
461 GGCGCTGATTAATATGCTGCTCGATCAATATATCCCTGCTCAATGAC 510
99 LysSerThrProAsnLeuValAArgSerTyrGlyPheTyrCysCysP 115
|||||
511 CAGGCAACACCCCGCATGACTCGGTGCTTACGAGTTTACGTTTACGCC 560
115 oGly...ThrGluGlyGluTyrCysGlyGlySerGlyGluSerPheC 131
|||||
561 AGGACCCCAATATGAAGAAATATGTGTAATATCCTCAGATTTCTTTT 610

```



```

131 ysaArgArgTTPSerCysValThrSerAsnAspGlyAspTTPlysTTPPro 147
115:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
611 GCACGCAATGAGAGCTGCATACCTCTTAATGATGGGAAATGGACCA 660
148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
661 GTCCTCAGCAACAGACAGAGTAAGTACTCTTTGTTAACAAATCCACACG 710
162 .....GlyProGlyLysTTPlys.....MetMetL 170
711 TTATAATCAATTTAATTAATGATGGGAGATGGAAGAAATGGCAACAC 760
170 ysaLeuTTPlysAspLys.....SerCysSerProSerAsp 181
761 GGGTACAAAAGATGTACGAATAAACAATAGCTGCATCTGTTAGAC 810
182 LeuAspTTPLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyr 197
811 CTAGATTACTTAATAAATAAGTTCTACTGAAAAAGAAAAACAAGAAATAT 860
198 SerLysValAspLysTTPArgLysLeuGlyAsnSer..PheLeuLeuTyrG 214
861 TCAAAAAGGGGTAAATGTATATCTTGGGGAATAGTATAC.....TATG 904
214 yGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgTle 227
905 GAGGCTCTGGGAGAAAGAGATCTGTCTGACTATTCGCCTCAGATA 954
228 GluTTPArgLysThrGluProProValAlaMetGlyProAspLysValLeuAl 244
955 GAACACTCAGATGGAACTCCGGTGTCTAATAGCAACCAATTAAGGGTTGGC 1004
244 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValProG 261
1005 CCAACAGAGACCTCCATCCAAAGACAG..... 1032
261 InLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 277
1033 .....AGGCCATCTCCTCAACCCCTGATTACATACATAACAC 1068
278 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValAlaPro 294
1069 TCTGGATAGTCCCACT.....GAGCCTTAACATCACTAT 1103
294 IlystThrGlyLysArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 311
1104 TAAACAGAGGGGAAACTTTTAGCTCATCCAGAGGAGCTTTCAAGCTC 1153
311 leaAsnSerThrAspProAspAlaThrSerSerCysTTPLeuCysLeuSer 327
1154 TTAACCTCCAGCACTCCAGAGGCTACTCTTCTGTGGCTTGGCTTAGCT 1203
328 SerGlyProProTyrTTPArgLysGlyMetAlaLysGlyLysPheAsnVa 344
1204 TCGGGCCCACTTACTATGAGGAAATGGCTAGAGGAGGAAATTCATGT 1253
344 IThrLysGlnHisArgAsnGlnCysThrTTPGlySerArgAsnLysLeu 361
1254 GACAAAGAACATAGAGACCAATGTACATGGGATCCCAAAATATAGCTTA 1303
361 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 377
1304 CCTTACTAGAGTTTCTGGAAAAAGCACTGCATAGGAGATGGTTCCCA 1353
378 SerHisGlnHisLeuCysTTPSerThrValValTyrGlnGlnAlaSerG 394
1354 TCCACCCACACCTTTGTAAACACACGTAAGCCTTTATATGCAACCTTGA 1403
394 uAsnGlnTTPLeuValProGlyTyrAsnArgTTPTrpAlaCysAsnThrG 411
1404 GAGTCATATATCTGCTACCTGTTATGACAGGTGTGGGCAATGTAATACTG 1453
411 LysLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427

```

```

1454 GATTAAACCCCTGTGTGTTCCACCTGTGTTTCAACCAACTAAAGACCTT 1503
428 CysValMetValGlnIleValProArgValTyrTTPArgProGlnGluVa 444
1504 TGCCTTATGTCCAAAATTTGCCCGGGGTGTACTACTATATCCGAAAAAGC 1553
444 IValLeuAspGlnTyrAspTTPArgTyrAsnArgProLysArgLysProV 461
1554 AGTCCTTGATATATATGACTATACATATATATGCGCCAAAAGAGAGCCCA 1603
461 aSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 477
1604 TATCCCTGACACCTAGAGCTGATATGCTCGATTGGGAGTGGCGCAGGGGTG 1653
478 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlySGI 494
1654 GGAACAGGAACGGCTGCCCTAATCAGAGACCCGCAACAGCTGGAGAAAG 1703
494 yLeuGlyGlnLeuHisAlaIleMetThrGluAspLeuArgAlaLeuLysG 511
1704 ACTTAGTAACCTTACATGCAATTGTACGAGAAATCTCCAAAGCCTAGAAA 1753
511 IuSerValSerAsnLeuGlnGluSerLeuThrSerLeuSerGluValVal 527
1754 AATCTGCAGTAACCTGAGAGAAATCCCTAACCTCTTATGTAAGTGT 1803
528 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlnGlyLe 544
1804 CTACAGAAACAGAAAGGGGTTAGATGTTATTTCTAAAGAGAGAGGTT 1853
544 uCysAlaAlaLeuLysGlnGluCysCysPheTyrValAspHisSerGly 561
1854 ATGTGTAGCCTTAAAGAGCAATGCTCTCTTATGTAATGATCAGCAGAG 1903
561 IalIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArg 577
1904 CCATCAGAGACTCCATGAGCAAGCTTAGAGAAAGTTAGAGAGCGCTGCA 1953
578 ArgGlnArgGlnAlaAspGlnGlyTTPPheGlnGlyTTPPheAsnArg 594
1954 AGGGAAGAGAGGCTGCACAGGGGTGTGTTGAAGAGATGTTCAACAGGTC 2003
594 rProTTPMetThrThrLeuLeuSerAlaLeuThrGlnGlyProLeuValVal 611
2004 TCTCTGGATACCACTGCTTCTGCTCTGACGGGGGCCCTTAGTAGTCTC 2053
611 euleLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 627
2054 TGCCTCGTTACTTACAGTTGGGCTTGCCTTATTAATAGCTTGTGTGCC 2103
628 PheValArgGlnArgValSerAlaValGlnIleMetValLeuArgGlnG 644
2104 TTTGTTAGAGAACAGATGAGTGCAGTCCAGATCATGTGTTAGGCACAA 2153
644 nTyrGlnGlyLeuLeuSerGlnGlyGlnTTPArgAspLeu 656
2154 GTACCAAGGCTTCTGACGCCAAGAGAAACTGACCTTC 2190
seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67022
seq_documentation_block:
ID AAC67022 standard; DNA: 4918 bp.
XX AAC67022:
XX
XX 27-MAR-2001 (first entry)
XX
XX PERV env protein coding sequence SEQ ID NO: 22.
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
XX Porcine endogenous retrovirus.

```

XX WO200071726-A1.  
 XX 30-NOV-2000.  
 XX 24-MAY-2000; 2000MO-US14296.  
 XX 24-MAY-1999; 9905-0135631.  
 XX (MAYO-) MAYO MEDICAL VENTURES.  
 XX Federspiel MJ;  
 XX WPI; 2001-032041/04.  
 XX  
 PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 PS Claim 16; Page 109-111; 144pp; English.  
 XX  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

alignment\_scores:  
 Quality: 2509.50 Length: 670  
 Ratio: 4.357 Gaps: 11  
 Percent Similarity: 85.970 Percent Identity: 70.896

alignment\_block:  
 US-09-171-553b-6 x AAC67022 ..

Align seg 1/1 to: AAC67022 from: 1 to: 4918

1 MethSerProThrLeuSerArgArgHisLeuProThrArgGlyGlyLeuPro 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2187 ATGCATCCACGCTTAAGCCGGCGACCTCCCGATTCGGGGTGGAAACCC 2236  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 17 OlySerArgLeuArgLeuProLeuSerPheLeuSerLeuLeuAlaTrpPheLeu 34  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2237 GAAAGAGCTGAAATCCCTTAAGCTTCGCTCCATCGCGTGGTCTCTTA 2286  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2287 CTCTGTCAATACCTCCAGTTAATGTTAAAGCGCTGTGGAGACGCCG 2336  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 51 AsnProHisArgProLeuSerLeuThrTrpLeuIleLeuAspProAsp 67  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2337 AACCTCCATTAACCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 2386  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2387 AGGTATTAATATTAACAGACATCAAGGGAGGCTCCCTGGGACCTGTGT 2436  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2437 GCCCTGAATTAATGTCTGCTTCATCATGAATCCCTGGTCTCAATGAC 2486  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 99 LysSerThrProProAsnLeuValArgSerTrpGlyPheTrpCysPr 115  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2487 CAGGCCACACCCCGGATGTACTCGTCTTAAGGGTTTACGTTTGCC 2536  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 115 oGly...ThrGlnLysGlnLysTrpCysGlyGlySerGlyLysLeuSerPheC 131  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2537 AGGACCCCAATATATGAAGATATTTGGAAATCCTCAGATTTTCTTTT 2586  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2587 GCAAGCAATGAGACCTGCTTAATGATGCAATGGAATGGAATGGCCA 2636  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2637 GTCTCTGACGACAGACAGTAAGTACTCTTTGTTACAAATCCTACCG 2686  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 162 .....GlyProGlyLysTrpLys.....MetMetL 170  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2687 TTATATCATATTAATTATGGCCATGGAGATGGAATGGCAACACG 2736  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2737 GGGTACAAAAGATGTACGAAATACCAATAATGCTGCTATTCGTTAGC 2786  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 182 LeuAspTrpLeuLysIleSerPheThrGln..ArgLysThrGlyLysTrp 197  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2787 CTAGATTACTTAAATAATAGTTTCACTGAAAAGAAAACAGCAAAATAT 2836  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 198 SerLysValAspLysTrpTrpGlnLeuGlyAsnSer..PheLeuLeuTrpG 214  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2837 TCAAAAGTGGGTAAATGATGTCTTGGGGAATAGTGTAC.....TATG 2880  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 214 LysGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgGly 227  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2881 GAGGCTCTGGGAGAAAGAAAGATCTGTTCTGACTATTCGCTCCACAGAAA 2930  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 228 GlnThrGlyThrGlnProProValAlaMetGlyProAspLysValLeuAl 244  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2931 GAAACTCAGATGGAACCTCCGGTTGCTATAGACCAATAAGGTTTGGC 2980  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 244 agLingLysProProAlaLeuGlnProProHisAsnLeuProValProG 261  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 2981 CGAAACAAGGACCTCCATCCAAAGACAG..... 3008  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 261 hLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 277  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3009 .....AGGCCATCTCTTACCCCTCTGATTAACATTAACACACC 3044  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 278 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 294  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3045 TCTGGATCATCCCACT.....GACCTTAACATCATCTAT 3079  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 294 LysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAla 311  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3080 TAAACAGGGGGCGAACTTTTAACCTCAATCCAGGAGCTTTTCAAGCTC 3129  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 311 LeuAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 327  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3130 TTAACTCCACGACTCCAGAGGCTACCTCTTGTGGCTTGTCTTAGCT 3179  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 328 SerGlyProProTrpTrpGlnGlyMetAlaLysGlnArgLysPheAsnVa 344  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3180 TCGGGCCACCACTTACTATGAGGAATGGCAGAGGGAATTAATCATATG 3229  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 344 LThrLysGlnHisArgAsnGlnCysThrTrpGlySerAlaGlnLysLeu 361  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3230 GACAAAGAACATAGAGACCAATGATGATGATGATGATGATGATGATGAT 3279  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 361 hLeuThrGlnValSerGlyGlyGlyThrCysTrpIleGlyLysAlaProPro 377  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3280 CCTTACTAGAGTTTCTGGAAAGGACCTGCTATGAGGATGGTTCGCCCA 3329  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 378 SerHisGlnHisLeuCysTrpSerThrValAlaTrpGlnGlnAlaSerG 394  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3330 TCCGACCAACACCTTTGTATACACACAGAACCTTAATACGAACTCTGA 3379  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 394 uAsnGlnTrpLeuValProGlyLysArgTrpTrpAlaCysAsnTrpG 411  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 3380 GAGTCAATATCTGTACTGTGTATGACAGGTGTGGGCTGTAAATACATG 3429  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 411 LysLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

```

|||||
3430 GATTAAACCCCTGTTGTTCCACCTTGTTTTCACCAACTAAGACTTT 3479
428 CysValMetValGlnIleValProArgValTyrTyrHisProGluGluVal 444
|||||
3480 TGGGTATGGTCCAAATGTGCCCCGGGGTGTACTACTATCCGGAAGAAC 3529
444 ValLeuAspGluTyrAspTyrArgTyrAsnArgProGlySerGluProVal 461
|||||
3530 AGTCCTTGATGATATGACTATGATATATGCGCCCAAAAGAGAGCCCA 3579
461 aIseIleuThrIleuValMetLeuGlyLeuGlyThrIleValGlyVal 477
|||||
3580 TATCCCTGACACTAGCTGTATAGCTCGATGGAGTGGCTGCGAGCGTG 3629
478 GlyThrGlyThrAlaIleuIleThrGlyProGlnGluLeuGluVal 494
|||||
3630 GGAAACAGAAAGCGCTGCCCTTAATCACAGGACCCCAACAGCTGAGAAAG 3679
494 yLeuGlyGluLeuHisAlaIleMetThrGluAspLeuArgIleGluVal 511
|||||
3680 ACTTAGTAACCTACATCGAATGTGTACGAGAGATCTCCAAAGCCCTAGAAA 3729
511 IuSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 527
|||||
3730 AATCTGCACTAACCTGAGAGATCCCTAACCTCTTATCTGAGAGTGTT 3779
528 LeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGluGlyVal 544
|||||
3780 CTACAGACAGAAAGGGGTAGATCTGTATTCTTAAGAAAGAGAGATT 3829
544 uCysAlaIleLeuGlyGluGlyCysPheTyrValAspHisSerGlyVal 561
|||||
3830 ATGTGTAGCTTGAGAGAGAAATGCTGTTTATGTGATCATTCAGGG 3879
561 IaIleArgAspSerMetAsnIleLeuArgLysLysLeuGluArgArgArg 577
3880 CCATCAGAGACTCCATACACAACTTAGAGAAAGTTGAGAGAGCGTCCA 3929
578 ArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSe 594
|||||
3930 AGGGAAGAAAGAACTACTCAAGGTGGTTGAGGAGTGTCAACAGGTC 3979
594 rProTyrMetThrThrIleuSerAlaLeuThrGlyProLeuValVal 611
|||||
3980 TCCTTGGTGGTACCTACTCTTCTGTATACAGGACCCCTTAATAGTCC 4029
611 euLeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgPheValAla 627
|||||
4030 TCCCTCCGTACTACAGAGTTGGCCATGTATTTATTAACAAGTTAATTGCC 4079
628 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnI 644
|||||
4080 TTCATTTAGAGACGAATAGTGCAGTCCAGATCATGTACTTAGACAACA 4129
644 nTyrGln 646
4130 GTACCAA 4136
seq_name: /SID2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC67018
seq_documentation_block:
ID AAC67018 standard; DNA; 1980 BP.
XX
XX AAC67018;
XX
XX 27-MAR-2001 (first entry)
XX
XX PERV-1-15 env protein coding sequence SEQ ID NO: 18.
XX
XX Xenotransplantation; infectious agent; vaccine; ds.
XX
XX Porcine endogenous retrovirus.

```

```

XX
XX NM0200071726-A1.
XX
XX 30-NOV-2000.
XX
XX 24-MAY-2000; 2000MO-0U514296.
XX
XX 24-MAY-1999; 99US-0135631.
XX
XX (MAYO-) MAYO MEDICAL VENTURES.
XX
XX Federspiel MJ;
XX
XX WPI: 2001-032041/04.
XX
XX Inhibiting or preventing infectious agent transmission in mammalian
XX transplant recipients, by introducing recombinant DNA comprising DNA
XX PT encoding extracellular proteins of the agent into donor cells, such as
XX PT swine cells.
XX
XX PS Claim 16; Page 100-101; 144pp; English.
XX
XX CC The present invention provides a method to prevent the transmission of
XX CC infectious agents during xenotransplantation. This involves introducing
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from
XX CC the infectious agent, and then introducing these cells into the
XX CC transplant recipient.
XX
XX SQ Sequence 1980 BP; 582 A; 452 C; 447 G; 499 T; 0 other;

```

```

alignment_scores:
    Quality: 2497.50      Length: 669
    Ratio: 4.343          Gaps: 11
    Percent Similarity: 85.949    Percent Identity: 70.852
alignment_block:
US-09-171-553b-6 x AAC67018 ..

```

Align seg 1/1 to: AAC67018 from: 1 to: 1980

```

1 MethIspProThrLeuSerArgArgHisLeuProThrArgGlyGluPr 17
|||||
1 ATGCATCCACGCTTAAGCGGGCCACCTCCGATTCGGGGTGAAGACC 50
17 oLysArgLeuArgGlyIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34
|||||
51 GAAAGAGACTGAATAATCCCTTAAGCTTGCCCTCCATCGCGTGTCCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
101 CTCTGTCAATTAACCTCTCAAGTTAATGTAAGCGCTGTGACAGCCG 150
51 AsnProHisArgProLeuSerLeuThrTyrLeuIleIleAspProAspPh 67
|||||
151 AACTCCCATTAACCCCTTATCTCTCACTGTACTACTAGTACTCGGATC 200
67 rGlyValIleThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||
201 AGGTATTAATTAATTAACAGCACTCAAGGGAGGAGCTCCCTTGGGAGCTGT 250
84 rPProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
|||||
251 GGCCTAATTAATATATGCTGCTCCATCAATGATATCCCTGTCTCAATGAC 300
99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
|||||
301 CAGGCGACACCCCGGATGCTACTCCGCTTACGGGATTTTACGTTTGCC 350
115 oGly...ThrGluGlyGluLysTyrCysGlyGlySerGlyIuSerPheC 131
|||||
351 AGGCGCCCAATTAATGAAGAAATTTGTGGAATCTCTCAGGATTTCTTTT 400

```

131 ysArgArgTTPSerCysValThrSerAsnAspGlyAspTTPPro 147  
 401 GCAAGCAATGAGCTGGTAACTTCAATGATGGAAATGGAAATGCCA 450  
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161  
 451 GTCTCTCAGCAAGACAGAGTAAGTACTCTTTGTTAACATCTACACAG 500  
 162 ..... GlyProGlyLysTyrLys..... MetMetL 170  
 501 TTATATCAATTTAATATGCGCATGGAGATGGAAAGATGGCAACAC 550  
 170 ysLeuTyrLysAspLys..... SerCysSerProSerAsp 181  
 551 GGGTACAAAAGATGTACCAATAAAGCAATAAGCTGATCTGTTAGAC 600  
 182 LeuAspTyrLeuLysLysLeuSerPheThrGlu.. ArgLysThrGlyLysTyr 197  
 601 CTAGATTACTTAATAAATAAGTTTCACTGAAAAAGAGAAACAAAGAAATAT 650  
 198 SerLysValAspLysTyrTyrGluLeuGlnLysSerPheLeu.. TyrG 214  
 651 TCAAAAGTGGTAATGATGATGTCTGGGGAATAGT... GlnCTATPAGAG 697  
 214 LysGlyLys..... AlaGlySerThrLeuThrIleArgLeuArgIleGlu 228  
 698 GCTCTGGGAGAAAGAAAGATGCTGTCTGACATTCGCTCGCATTAAGA 747  
 229 ThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaG1 245  
 748 ACTCAGATGAGACCTCCGCTGCTATAGACCAATAAGGGTTGGCCGA 797  
 245 uGlnGlyProProAlaLeuGluProProHisAsnLeuProValProGlnL 262  
 798 ACAAGGACTCCATCCAGAACAG..... 822  
 262 eutHrSerLeuArgProAspIleThrGlnProProSerAsnSerThrThr 278  
 823 ..... AGCCATCTCCTTACCCCTGATTACAAATACAAACCTCT 861  
 279 GlyLeuIleProThrAsnThrProArgAsnSerProGlyValProValLys 295  
 862 GGATCAGTCCCACT..... GAGCCTACATCACTATATTA 896  
 295 sThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleA 312  
 897 AACAGGGCGGAACCTTTTAACTCATCAGAGGAGCTTTCAACCTCTTA 946  
 312 snSerThrAspProAspAlaThrSerSerCysTyrLeuCysLeuSerSer 328  
 947 ACTCCAGACATCCAGAGGCTACCTCTTGTGCTTGTGCTTGTGCTTCG 996  
 329 GlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValTh 345  
 997 GGGCCACCTTAATATGAGGAGATGGCTAGAGAGGAAATTCATATGTGAC 1046  
 345 rLysGlnHisArgAsnGlnCysThrTyrPglYserArgAsnLysLeuThrL 362  
 1047 AAAAGAACATAGAGACCAATGTACATGGGATCCCAAAATAGCTTACCC 1096  
 362 eutHrGluValSerGlyLysGlyThrCysIleGlyLysAlaProProSer 378  
 1097 TTACTGAGGTTTCTGAAAAGGACCTGCATAGGAGATGTTCCGCCATCC 1146  
 379 HisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGluAs 395  
 1147 CACCAACACCTTGTATACACACATGAGACCTTAATAGAACCTGTAGAG 1196  
 395 nGlnTyrLeuValProGlyTyrAsnArgTTPTPalacysAsnThrGlyL 412  
 1197 TCACATATCTGCTACTGCTGATGACAGTGTGGGCTGTAATACCTGAT 1246  
 412 eutHrProCysValSerThrSerValPheAsnGlnSerLysAspPheCys 428

1247 TACCCCTGTGTTCACACCTTGTTTCAACAAACAAAGACTTTTGC 1296  
 429 ValMetValGlnIleValAlaProArgValTyrTyrHisProGluGluValVa 445  
 1297 GTTATGGTCCAAATTTGTCCTCCGGGTGTACTACTATCCGAAAAAGCACT 1346  
 445 IleuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProValS 462  
 1347 CCTTGATGATATGATGATATGATATTAATCCGCAAAAAAGAGGCCATAT 1396  
 462 erLeuThrLeuAlaValMetLeuGlnLysGlyThrAlaValAlaGlyValGly 478  
 1397 CCTGACACATGCTGTATGCTCGGATTTGGAGTGGCTGCAAGCGCTGGA 1446  
 479 ThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlnLysGlyLe 495  
 1447 ACAGGAAGCGCTGCCCTTAATCAGAGAACCCGCAACAGCTGGAGAAAGACT 1496  
 495 uGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluS 512  
 1497 TAGTAACTCATATCAATTAACGGAATAATTCACAGCCCTAGAAAAAT 1546  
 512 erValSerAsnLeuGlnGluSerLeuThrSerLeuSerGluValAlaLeu 528  
 1547 CTGTACAGTAACTGAGAGAAATCCCTTAACCTCTTATCTGGAAGGTTCTA 1596  
 529 GlnAsnArgArgGlyLeuAspLeuPheLeuArgGluGlnGlyLeuLys 545  
 1597 CAGAGCAAGAGGGGTAGATCTGTATTTCTTAAAGAGAGAGATTATG 1646  
 545 sAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaI 562  
 1647 TGTAGCCTTAAGAGAGAAATGCTGTTTATGTGATCATTTAGGGGCCA 1696  
 562 leuArgAspSerMetAsnLysLeuArgLysLysLeuGluArgAlaGlyArg 578  
 1697 TCAGAGCTCTCATAGAACCAAGCTTAGAGAAAGTGGAGAACCTCGAAGG 1746  
 579 GluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSerP 595  
 1747 GAAAGGAACACTACTCAGAGGTGCTTGTGAGGATGTTTAAACAGCTTCC 1796  
 595 cTTPMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuL 612  
 1797 TTGTTGGCTACCTCTTCTCTTCTTAAACAGAGCCCTTAATAGCTTCC 1846  
 612 euleuLeuThrValGlyProCysLeuLLeuAsnArgPheValAlaPhe 628  
 1847 TCTGTACTCACAGTTGGGCAATGATATTAACAAAGTTAATGCTTTC 1896  
 629 ValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnT 645  
 1897 ATTAGAGAAAGAAATAGTGCAGATCCAGATCATGACTTACCAACACAGTA 1946  
 645 rGln 646  
 1947 CCNA 1950

seq\_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAF77726

seq\_documentation\_block: ID AAF77726 standard; DNA: 7333 BP.

AAF77726;

23-MAY-2001 (first entry)

Defective retroviral genome isolated from PK-15 cell line.

Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;

ss.

OS Unidentified.  
 XX US6190861-B1.  
 PN 20-FEB-2001.  
 PD 13-DEC-1996: 9605-0766528.  
 PF 14-DEC-1995: 9505-0572645.  
 PR (GEHO ) GEN HOSPITAL CORP.  
 PA Fishman JA;  
 XX WPI: 2001-256211/26.  
 DR P-P-SDB: AAB73282, AAB73283, AAB73284.  
 XX Assessing risk of endogenous retroviruses in clinical practice and in  
 PT xenotransplantation, comprises using probe sequences derived from swine  
 PT or miniature swine retroviral genome -  
 XX Claim 1; Fig 2; 127pp; English.  
 XX The present invention relates to a method for screening a cell or tissue  
 CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g. the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.  
 XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;  
 SQ  
 alignment\_scores:  
 Quality: 2497.50 Length: 678  
 Ratio: 4.328 Gaps: 13  
 Percent Similarity: 85.103 Percent Identity: 71.829  
 alignment\_block:  
 US-09-171-553B-6 x AAF77726 ..  
 Align seg 1/1 to: AAF77726 from: 1 to: 7333  
 4 ThrLeuSerArgArgHisLeuProThrArgGlyGluProLysArgLe 20  
 |||||  
 4757 ACCCTATCTCGTACTTTGACCAACGCGCTGTGAAGTCGA..... 4801  
 20 uArg illeProLeuSerPheAlaSerIleAlaTrpPheLeuThr 36  
 |||||  
 4802 .AGGAATCCCTTAAGCTTCGCTCCATCGCGTGGTCTTACTACTGTC 4850  
 37 IleThrProGlnAlaSerSerLysArgLeuLeuSerSerSerAla 53  
 |||||  
 4851 ATAACTCTCAAGTTAATGTAAGCGCTTGTGGACACCCGGAACCTCCA 4900  
 53 sArgProLeuSerLeuThrTrpLeuIleLeuSerProAspThrGlyAl 70  
 |||||  
 4901 TAAACCTATCTCGTACCTGGTCTACTACTGACCGGTACAGGTATTA 4950  
 70 hValAsnSerThrArgGlyAlaIleProArgGlyThrTrpProLys 86  
 |||||  
 4951 AATATTAAACAGCACTCAAGGAGGCTCCCTGGGAGCTGTGGCTTAA 5000  
 87 LeuHisPheCysLeuArgLeuLeuAsnProAlaVal.....LysSerTh 101  
 |||||  
 5001 TTATATGTGTGCTTCGATCAGATATCCGTGTCTCATATGACACAGCC 5050  
 101 rProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGly...T 117

|||||  
 5051 ACCCCCGATGTACTCCGTGCTTACGGGTTTACGTTGCCAGGACCCC 5100  
 117 hrcLulysglulysTyrCysGlySerGlyGluSerPheCysArg 133  
 |||||  
 5101 CAATATATGAGAAATATGTGGAAATCTCAGGATTTCTTTGGACAGCAA 5150  
 134 TrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 150  
 |||||  
 5151 TCGAGCTGCATTACTTCTTAATGATGGGAATGTGAATGGCCAGTCTCTCA 5200  
 150 uGlnAspArgValLysPheSerPheValAsnSer..... 161  
 |||||  
 5201 GCAAGACAGAGTAAGTTACTCTTTGTAAACATCTACAGTTATATATC 5250  
 162 .....GlyProGlyLysTyrLys.....MethLysLeuTyr 172  
 |||||  
 5251 AATTATATTGTGGCATGGGATGGAAGATTGGCAACAGCGGGTACAA 5300  
 173 LysAspLys.....SerCysSerProSerAspLeuAspTyr 184  
 |||||  
 5301 AAGATGTACGAATAAAGCAATAGCTGTCAATTCGTTAGACCTAGATTAA 5350  
 184 rLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyrSerLys 200  
 |||||  
 5351 CTTAAAAATTAAGTTTCACTGMAAAAGAAACAAAGAAATATTCAAAGT 5400  
 200 LAspLysTrpTyrGluLeuGlyAsnSer..PheLeuLeuTyrGlyGly 216  
 |||||  
 5401 GGGTAATAGTATATCTTGGGAAATAGTGTAC.....TATGAGGCTCT 5444  
 217 Ala.....GlySerThrLeuThrIleArgLeuArgIleGluThr 230  
 |||||  
 5445 GGGAGAAAGAAAGATCTCTTGTGACTATTCGCTCAGAAATAGAACTCA 5494  
 5495 GATGGAACTCCGCTTGTAGTACCAATAAGAGGCTTGGCCGACACAG 5544  
 247 LyrProProAlaLeuGluProProHisAsnLeuProValProGlnLeu 263  
 |||||  
 5545 GACCTCCAAATCCAAAGAACAG..... 5564  
 264 SerLeuArgProAspIleThrGlnProProSerAsnSerThrThrGly 280  
 |||||  
 5565 .....AGGCATCTCTTAACCCCTCGATTACATACAACTCTGGATC 5608  
 280 uIleProThrAsnThrProArgAsnSerProGlyValProValLysThr 297  
 |||||  
 5609 AGTCCCACT.....GAGCCTACATCACTATTAAACAG 5643  
 297 LysIleArgLeuPheSerLeuIleGlnIleValIlePheGlnAlaIleAsn 313  
 |||||  
 5644 GGGGAAACTTTTACCTCATTCACAGGAGCTTTTCAAGCTCTTAATCTC 5693  
 314 ThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGly 330  
 |||||  
 5694 ACGACTCCAGAGGCTACCTCTTCTTGTGGCTTGTACTTGGGCCC 5743  
 330 oProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLys 347  
 |||||  
 5744 ACCTTACTATGAGGAATGCTAGAGGAGGAATTAAGTGAACAAAG 5793  
 347 IuHisArgAsnGlnCysThrTrpLysSerArgAsnLysLeuThrLeu 363  
 |||||  
 5794 AACATAGACCAATGTACATGGGATCCCAAAATAAGCTTACCTTACT 5843  
 364 GluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisG 380  
 |||||  
 5844 GAGGTTCTGTGAAAAGGACCTGCATAGGATGATGTCCTCCCATGCCCA 5893  
 380 nHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGluAsnGln 397  
 |||||

```

5894 ACAACCTTTGTACACACAGTGAACCTTTAATGACACCTCTGACAGTCATC 5943
397 YLLeuValProGlyTYrAsnArgTTPAlaCysAsnThrGlyLeuThr 413
|||||
5944 ATCTGTAACCTGGTTATGACAGGTGGCGATGATATCTGATTAAC 5993
414 ProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMet 430
|||||
5994 CCTTGTTTCCACCTTGGTTTTCACCAACTAAAGACTTTTGCTTAT 6043
430 tValGlnIleValProArgValIleTYrHisProGlnIleValIleVal 447
|||||
6044 GGTCCAAATTTGCCCGGGGTGACTACTATCCCGAAAGACAGTCCCTTG 6093
447 spGluTYrAspTYrArgTYrAsnArgProLysArgGluProValSerLeu 463
|||||
6094 ATGAATATGACTATAGATATAAATCGCCAAAAGAGAGCCCATATCCCTG 6143
464 ThrLeuAlaValMetLeuGlyLeuGlyThrAlaValAlaGlyIleThrGI 480
|||||
6144 ACACGTAGCTGTAATGCTCGGATTGGAGAGGCTGCACGGCTGGGAAACAG 6193
480 YThrAlaAlaLeuIleThrGlyProGlnIleLeuGlnLysGlyLeuGlyG 497
|||||
6194 AACGGCTGCCCTAATCACAGACCCCAACAGCTGCAGAAAGACACTAGTA 6243
497 ILeuHisAlaIleMetThrGluAspLeuArgAlaLeuLysGluSerVal 513
|||||
6244 ACCTACATCGAATTTGTAAAGGAGATCTCAAGCCCTAGAAAACCTCTGC 6293
514 SerAsnLeuGluGluSerLeuThrSerLeuSerGluValIleLeuGlnAs 530
|||||
6294 ACTAACCTGGAGAGATCCCTAACCTCTTATCTGTAAGTGGTTCTACAGAA 6343
530 nArgArgGlyLeuAspLeuPheLeuArgGluGlyLeuGlyLeuGlyLeuG 547
|||||
6344 CAGAAAGGGGTTAGATCTGTTATTTCTAAAGAGAGAGGTTATGCTAG 6393
547 IAlaLeuLysGluGluCysCysPheTYrValIleAspHisSerGlyAlaIle 563
|||||
6394 CCTTAAAAAGAGAAATGCTGCTCTTCTATGTAGATCACTCAGAGCCATCA 6443
564 AspSerMetAsnLysLeuArgLysLysLeuGluArgArgArgGluArg 580
|||||
6444 GACTCATGAGCAACCTTAGAGAAAGTTAGAGAGCCCTGCAAGGGAAG 6493
580 GGLuAlaSpGlnGlyTTPheGlnGlyTTPheAsnArgSerProTirPM 597
|||||
6494 AGAGGCTGACCAAGGGGTGTTGAAGGATGCTTCAACAGGCTCTCTTGA 6543
597 eThrThrLeuLeuSerAlaLeuThrGlyProLeuValIleValLeuLeu 613
|||||
6544 TGACCAACCTGCTTCTCTCTGACAGGGGCCCTAGTAGTCTCTCTCTG 6593
614 LeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheVal 630
|||||
6594 TTACTTACAGTTGGGCTTCTCTTAAATAGGTTTGTGCTTGTAG 6643
630 GGLuArgValSerAlaValGlnIleMetValLeuArgGlnIleTYrGln 647
|||||
6644 AGAAGAGGTGAGTGCAGTCCAGATCAGTACGTAAGCAACAGTACCAAG 6693
647 ILeuLeuSerGlnGlyGluThrAspLeu 656
|||||
6694 GCCTTCGAGCCAGAGAAACTACCTC 6722

```

seq\_name: /STDSD/gcgdata/geneseq/geneseq/NA1997.DAT: AAT74883

seq\_documentation\_block:

ID AAT74883 standard: cDNA: 7393 BP.

XX AC AAT74883;

XX

```

DT 09-FEB-1998 (first entry)
XX Porcine retrovirus cDNA (defective).
DE Retrovirus: porcine; GAG protein; POL protein; ENV protein;
XX xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; PCR; ss.
XX Porcine retrovirus.
OS
XX
FH Key Location/Qualifiers
FT CDS 598..2172
FT /tag= a
FT /tag= b
FT /tag= c
FT /tag= d
FT /tag= e
FT /tag= f
FT /tag= g
FT /tag= h
FT /tag= i
FT /tag= j
FT /tag= k
FT /tag= l
FT /tag= m
FT /tag= n
FT /tag= o
FT /tag= p
FT /tag= q
FT /tag= r
FT /tag= s
FT /tag= t
FT /tag= u
FT /tag= v
FT /tag= w
FT /tag= x
FT /tag= y
FT /tag= z
FT /tag= aa
FT /tag= ab
FT /tag= ac
FT /tag= ad
FT /tag= ae
FT /tag= af
FT /tag= ag
FT /tag= ah
FT /tag= ai
FT /tag= aj
FT /tag= ak
FT /tag= al
FT /tag= am
FT /tag= an
FT /tag= ao
FT /tag= ap
FT /tag= aq
FT /tag= ar
FT /tag= as
FT /tag= at
FT /tag= au
FT /tag= av
FT /tag= aw
FT /tag= ax
FT /tag= ay
FT /tag= az
FT /tag= ba
FT /tag= bb
FT /tag= bc
FT /tag= bd
FT /tag= be
FT /tag= bf
FT /tag= bg
FT /tag= bh
FT /tag= bi
FT /tag= bj
FT /tag= bk
FT /tag= bl
FT /tag= bm
FT /tag= bn
FT /tag= bo
FT /tag= bp
FT /tag= bq
FT /tag= br
FT /tag= bs
FT /tag= bt
FT /tag= bu
FT /tag= bv
FT /tag= bw
FT /tag= bx
FT /tag= by
FT /tag= bz
FT /tag= ca
FT /tag= cb
FT /tag= cc
FT /tag= cd
FT /tag= ce
FT /tag= cf
FT /tag= cg
FT /tag= ch
FT /tag= ci
FT /tag= cj
FT /tag= ck
FT /tag= cl
FT /tag= cm
FT /tag= cn
FT /tag= co
FT /tag= cp
FT /tag= cq
FT /tag= cr
FT /tag= cs
FT /tag= ct
FT /tag= cu
FT /tag= cv
FT /tag= cw
FT /tag= cx
FT /tag= cy
FT /tag= cz
FT /tag= da
FT /tag= db
FT /tag= dc
FT /tag= dd
FT /tag= de
FT /tag= df
FT /tag= dg
FT /tag= dh
FT /tag= di
FT /tag= dj
FT /tag= dk
FT /tag= dl
FT /tag= dm
FT /tag= dn
FT /tag= do
FT /tag= dp
FT /tag= dq
FT /tag= dr
FT /tag= ds
FT /tag= dt
FT /tag= du
FT /tag= dv
FT /tag= dw
FT /tag= dx
FT /tag= dy
FT /tag= dz
FT /tag= ea
FT /tag= eb
FT /tag= ec
FT /tag= ed
FT /tag= ee
FT /tag= ef
FT /tag= eg
FT /tag= eh
FT /tag= ei
FT /tag= ej
FT /tag= ek
FT /tag= el
FT /tag= em
FT /tag= en
FT /tag= eo
FT /tag= ep
FT /tag= eq
FT /tag= er
FT /tag= es
FT /tag= et
FT /tag= eu
FT /tag= ev
FT /tag= ew
FT /tag= ex
FT /tag= ey
FT /tag= ez
FT /tag= fa
FT /tag= fb
FT /tag= fc
FT /tag= fd
FT /tag= fe
FT /tag= ff
FT /tag= fg
FT /tag= fh
FT /tag= fi
FT /tag= fj
FT /tag= fk
FT /tag= fl
FT /tag= fm
FT /tag= fn
FT /tag= fo
FT /tag= fp
FT /tag= fq
FT /tag= fr
FT /tag= fs
FT /tag= ft
FT /tag= fu
FT /tag= fv
FT /tag= fw
FT /tag= fx
FT /tag= fy
FT /tag= fz
FT /tag= ga
FT /tag= gb
FT /tag= gc
FT /tag= gd
FT /tag= ge
FT /tag= gf
FT /tag= gg
FT /tag= gh
FT /tag= gi
FT /tag= gj
FT /tag= gk
FT /tag= gl
FT /tag= gm
FT /tag= gn
FT /tag= go
FT /tag= gp
FT /tag= gq
FT /tag= gr
FT /tag= gs
FT /tag= gt
FT /tag= gu
FT /tag= gv
FT /tag= gw
FT /tag= gx
FT /tag= gy
FT /tag= gz
FT /tag= ha
FT /tag= hb
FT /tag= hc
FT /tag= hd
FT /tag= he
FT /tag= hf
FT /tag= hg
FT /tag= hh
FT /tag= hi
FT /tag= hj
FT /tag= hk
FT /tag= hl
FT /tag= hm
FT /tag= hn
FT /tag= ho
FT /tag= hp
FT /tag= hq
FT /tag= hr
FT /tag= hs
FT /tag= ht
FT /tag= hu
FT /tag= hv
FT /tag= hw
FT /tag= hx
FT /tag= hy
FT /tag= hz
FT /tag= ia
FT /tag= ib
FT /tag= ic
FT /tag= id
FT /tag= ie
FT /tag= if
FT /tag= ig
FT /tag= ih
FT /tag= ii
FT /tag= ij
FT /tag= ik
FT /tag= il
FT /tag= im
FT /tag= in
FT /tag= io
FT /tag= ip
FT /tag= iq
FT /tag= ir
FT /tag= is
FT /tag= it
FT /tag= iu
FT /tag= iv
FT /tag= iw
FT /tag= ix
FT /tag= iy
FT /tag= iz
FT /tag= ja
FT /tag= jb
FT /tag= jc
FT /tag= jd
FT /tag= je
FT /tag= jf
FT /tag= jg
FT /tag= jh
FT /tag= ji
FT /tag= jj
FT /tag= jk
FT /tag= jl
FT /tag= jm
FT /tag= jn
FT /tag= jo
FT /tag= jp
FT /tag= jq
FT /tag= jr
FT /tag= js
FT /tag= jt
FT /tag= ju
FT /tag= jv
FT /tag= jw
FT /tag= jx
FT /tag= jy
FT /tag= jz
FT /tag= ka
FT /tag= kb
FT /tag= kc
FT /tag= kd
FT /tag= ke
FT /tag= kf
FT /tag= kg
FT /tag= kh
FT /tag= ki
FT /tag= kj
FT /tag= kk
FT /tag= kl
FT /tag= km
FT /tag= kn
FT /tag= ko
FT /tag= kp
FT /tag= kq
FT /tag= kr
FT /tag= ks
FT /tag= kt
FT /tag= ku
FT /tag= kv
FT /tag= kw
FT /tag= kx
FT /tag= ky
FT /tag= kz
FT /tag= la
FT /tag= lb
FT /tag= lc
FT /tag= ld
FT /tag= le
FT /tag= lf
FT /tag= lg
FT /tag= lh
FT /tag= li
FT /tag= lj
FT /tag= lk
FT /tag= lm
FT /tag= ln
FT /tag= lo
FT /tag= lp
FT /tag= lq
FT /tag= lr
FT /tag= ls
FT /tag= lt
FT /tag= lu
FT /tag= lv
FT /tag= lv
FT /tag= lw
FT /tag= lx
FT /tag= ly
FT /tag= lz
FT /tag= ma
FT /tag= mb
FT /tag= mc
FT /tag= md
FT /tag= me
FT /tag= mf
FT /tag= mg
FT /tag= mh
FT /tag= mi
FT /tag= mj
FT /tag= mk
FT /tag= ml
FT /tag= mm
FT /tag= mn
FT /tag= mo
FT /tag= mp
FT /tag= mq
FT /tag= mr
FT /tag= ms
FT /tag= mt
FT /tag= mu
FT /tag= mv
FT /tag= mw
FT /tag= mx
FT /tag= my
FT /tag= mz
FT /tag= na
FT /tag= nb
FT /tag= nc
FT /tag= nd
FT /tag= ne
FT /tag= nf
FT /tag= ng
FT /tag= nh
FT /tag= ni
FT /tag= nj
FT /tag= nk
FT /tag= nl
FT /tag= nm
FT /tag= nn
FT /tag= no
FT /tag= np
FT /tag= np
FT /tag= nq
FT /tag= nr
FT /tag= ns
FT /tag= nt
FT /tag= nu
FT /tag= nv
FT /tag= nw
FT /tag= nx
FT /tag= ny
FT /tag= nz
FT /tag= oa
FT /tag= ob
FT /tag= oc
FT /tag= od
FT /tag= oe
FT /tag= of
FT /tag= og
FT /tag= oh
FT /tag= oi
FT /tag= oj
FT /tag= ok
FT /tag= ol
FT /tag= om
FT /tag= on
FT /tag= oo
FT /tag= op
FT /tag= oq
FT /tag= or
FT /tag= os
FT /tag= ot
FT /tag= ou
FT /tag= ov
FT /tag= ow
FT /tag= ox
FT /tag= oy
FT /tag= oz
FT /tag= pa
FT /tag= pb
FT /tag= pc
FT /tag= pd
FT /tag= pe
FT /tag= pf
FT /tag= pg
FT /tag= ph
FT /tag= pi
FT /tag= pj
FT /tag= pk
FT /tag= pl
FT /tag= pm
FT /tag= pn
FT /tag= po
FT /tag= pp
FT /tag= pq
FT /tag= pr
FT /tag= ps
FT /tag= pt
FT /tag= pu
FT /tag= pv
FT /tag= pw
FT /tag= px
FT /tag= py
FT /tag= pz
FT /tag= qa
FT /tag= qb
FT /tag= qc
FT /tag= qd
FT /tag= qe
FT /tag= qf
FT /tag= qg
FT /tag= qh
FT /tag= qi
FT /tag= qj
FT /tag= qk
FT /tag= ql
FT /tag= qm
FT /tag= qn
FT /tag= qo
FT /tag= qp
FT /tag= qq
FT /tag= qr
FT /tag= qs
FT /tag= qt
FT /tag= qu
FT /tag= qv
FT /tag= qw
FT /tag= qx
FT /tag= qy
FT /tag= qz
FT /tag= ra
FT /tag= rb
FT /tag= rc
FT /tag= rd
FT /tag= re
FT /tag= rf
FT /tag= rg
FT /tag= rh
FT /tag= ri
FT /tag= rj
FT /tag= rk
FT /tag= rl
FT /tag= rm
FT /tag= rn
FT /tag= ro
FT /tag= rp
FT /tag= rq
FT /tag= rr
FT /tag= rs
FT /tag= rt
FT /tag= ru
FT /tag= rv
FT /tag= rw
FT /tag= rx
FT /tag= ry
FT /tag= rz
FT /tag= sa
FT /tag= sb
FT /tag= sc
FT /tag= sd
FT /tag= se
FT /tag= sf
FT /tag= sg
FT /tag= sh
FT /tag= si
FT /tag= sj
FT /tag= sk
FT /tag= sl
FT /tag= sm
FT /tag= sn
FT /tag= so
FT /tag= sp
FT /tag= sq
FT /tag= sr
FT /tag= ss
FT /tag= st
FT /tag= su
FT /tag= sv
FT /tag= sw
FT /tag= sx
FT /tag= sy
FT /tag= sz
FT /tag= ta
FT /tag= tb
FT /tag= tc
FT /tag= td
FT /tag= te
FT /tag= tf
FT /tag= tg
FT /tag= th
FT /tag= ti
FT /tag= tj
FT /tag= tk
FT /tag= tl
FT /tag= tm
FT /tag= tn
FT /tag= to
FT /tag= tp
FT /tag= tq
FT /tag= tr
FT /tag= ts
FT /tag= tt
FT /tag= tu
FT /tag= tv
FT /tag= tw
FT /tag= tx
FT /tag= ty
FT /tag= tz
FT /tag= ua
FT /tag= ub
FT /tag= uc
FT /tag= ud
FT /tag= ue
FT /tag= uf
FT /tag= ug
FT /tag= uh
FT /tag= ui
FT /tag= uj
FT /tag= uk
FT /tag= ul
FT /tag= um
FT /tag= un
FT /tag= uo
FT /tag= up
FT /tag= uq
FT /tag= ur
FT /tag= us
FT /tag= ut
FT /tag= uu
FT /tag= uv
FT /tag= uv
FT /tag= vw
FT /tag= vx
FT /tag= vy
FT /tag= vz
FT /tag= wa
FT /tag= wb
FT /tag= wc
FT /tag= wd
FT /tag= we
FT /tag= wf
FT /tag= wg
FT /tag= wh
FT /tag= wi
FT /tag= wj
FT /tag= wk
FT /tag= wl
FT /tag= wm
FT /tag= wn
FT /tag= wo
FT /tag= wp
FT /tag= wq
FT /tag= wr
FT /tag= ws
FT /tag= wt
FT /tag= wu
FT /tag= wv
FT /tag= ww
FT /tag= wx
FT /tag= wy
FT /tag= wz
FT /tag= xa
FT /tag= xb
FT /tag= xc
FT /tag= xd
FT /tag= xe
FT /tag= xf
FT /tag= xg
FT /tag= xh
FT /tag= xi
FT /tag= xj
FT /tag= xk
FT /tag= xl
FT /tag= xm
FT /tag= xn
FT /tag= xo
FT /tag= xp
FT /tag= xq
FT /tag= xr
FT /tag= xs
FT /tag= xt
FT /tag= xu
FT /tag= xv
FT /tag= xv
FT /tag= xw
FT /tag= xx
FT /tag= xy
FT /tag= xz
FT /tag= ya
FT /tag= yb
FT /tag= yc
FT /tag= yd
FT /tag= ye
FT /tag= yf
FT /tag= yg
FT /tag= yh
FT /tag= yi
FT /tag= yj
FT /tag= yk
FT /tag= yl
FT /tag= ym
FT /tag= yn
FT /tag= yo
FT /tag= yp
FT /tag= yq
FT /tag= yr
FT /tag= ys
FT /tag= yt
FT /tag= yu
FT /tag= yv
FT /tag= yw
FT /tag= yx
FT /tag= yy
FT /tag= yz
FT /tag= za
FT /tag= zb
FT /tag= zc
FT /tag= zd
FT /tag= ze
FT /tag= zf
FT /tag= zg
FT /tag= zh
FT /tag= zi
FT /tag= zj
FT /tag= zk
FT /tag= zl
FT /tag= zm
FT /tag= zn
FT /tag= zo
FT /tag= zp
FT /tag= zq
FT /tag= zr
FT /tag= zs
FT /tag= zt
FT /tag= zu
FT /tag= zv
FT /tag= zv
FT /tag= zw
FT /tag= zx
FT /tag= zy
FT /tag= zz

```

This cDNA sequence represents a defective purified swine retrovirus found in PK-15 cells containing the putative coding regions for viral GAG, POL, and ENV proteins. There are a few in frame stop codons and apparent frame shifts in the given coding sequence which alter features of the translation. This sequence and PCR fragments generated from the sequence (see AAT74812-T74882) could be used to screen organs for the presence of porcine retroviruses prior to xenotransplantation. CC Transplantation can increase the likelihood of retroviral activation if CC intact and infectious proviruses are present. The porcine retroviral CC sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus sequences in a strain of xenograft transplantation donors. It can be used to detect mutations, genetic lesions or viral recombinants and to determine the histological localisation of activated retroviruses. Using CC Polymerase Chain Reaction DNA quantitation (PQ) on blood mononuclear cells, infectivity titration and susceptibility testing can be

CC performed. Ultimately animal donors without intact porcine retroviral  
 CC sequences or a lower copy number of viral elements could be selected.  
 XX  
 SQ Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;

alignment\_scores:  
 Quality: 2497.50 Length: 678  
 Ratio: 4.328 Gaps: 13  
 Percent Similarity: 85.103 Percent Identity: 71.829

alignment block:

US-09-171-553b-6 x AAT74883 ..

Align seg 1/1 to: AAT74883 from: 1 to: 7393

```

4 ThrLeuSerArgArgHisLeuProThrArgGlyGlyProLysArgLe 20
|||||
4757 ACCTTATCTGCTACTTTTGACCCACACCAAGGCTGTGAAGTGA.... 4801
|||||
20 uArg, IleProLeuSerPheAlaSerIleAlaThrPheLeuThrLeuThr 36
|||||
4802 .AGGAATCCCTTAAGCTTCGCTCCATCCGCTGCTTCTACTCTGTCA 4850
|||||
37 IleThrProGlnAlaSerSerLysArgLeuIleAspSerSerAsnProH 53
|||||
4851 ATACCTCTCAAGTAAAGTAAACGCTTGAGACGCCGCAACCTCCCA 4900
|||||
53 sArgProLeuSerLeuThrTrpLeuIleLeuAspProAspThrGlyValT 70
|||||
4901 TAAACCTTATCTCTACCTGCTGTTACTTCTGACTCGGTACAGGATTTA 4950
|||||
70 hValAlaSerThrArgGlyValAlaProArgGlyThrTrpTrpProGln 86
|||||
4951 ATATATACAGCACTCAAGGAGGAGCTCCCTGGGAGCTGAGCCCTGAA 5000
|||||
87 LeuHisPheCysLeuArgLeuIleAsnProAlaVal..... LysSerTh 101
|||||
5001 TTATATGTCTGCTTCGATCATATATCCCTGCTTCATGACCAAGCCAC 5050
|||||
101 rProAsnLeuValArgSerThrGlyPheThrCysCysProGly...T 117
|||||
5051 ACCCCCGAGTACTCCGCTGCTTACGGGTTTACGTTGCCAGACCC 5100
|||||
117 hGlyLeuGlyLysThrCysGlyGlySerGlyLysPheCysArgArg 133
|||||
5101 CAAATATGAGAAATATGTGGAATCTCAGGATTTCTTTGCAAGCA 5150
|||||
134 TrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLe 150
|||||
5151 TGGAGCTGCATACCTTCTAATATGAGGAATGGGAATGGCCAGTCTCA 5200
|||||
150 uGlnAspArgValLysPheSerPheValAsnSer..... 161
|||||
5201 GGAAGACAGATAGTACTCTTTTGTTRACATCTCTACCACTTAAATC 5250
|||||
162 .....GlyProGlyLysThrLys.....MetLeuLysLeuThr 172
|||||
5251 AATTTAATTTATGCGCATGGAGATGGAAGATGGCAACAGGGGTACAA 5300
|||||
173 LysAspLys.....SerCysSerProSerAspLeuAspTyr 184
|||||
5301 AAAGATGTACGAATAATAGCAATAATAGCTGATTCGTAGACATGATTA 5350
|||||
184 rLeuLysIleSerPheThrGlu...ArgLysThrGlyLysThrSerLysVa 200
|||||
5351 CTTAAATAATAGTTTCACTGAAAAAGAAACAGAAATAATTCAAAAGT 5400
|||||
200 LAspLysTrpTyrGluLeuGlyAsnSer...PheLeuLeuThrGlyGly 216
|||||
5401 GGGTAAATGATATCTTGGGGAATAGTGTAC.....TATGAGAGGCTCT 5444
|||||
217 Ala.....GlySerThrLeuThrIleArgLeuArgIleGluThrGln 230

```

```

5445 GGGAGAAAGAAAGATGTGTCTGACTATGTGCTCAGAAATGAAACATCA 5494
|||||
230 yThrGluProProValAlaMetGlyProAspLysValLeuAlaGluGlnG 247
|||||
5495 GATGGAACTCCGGTTGCTATATAGGACAAATTAAGGTTTGGCCGAACAG 5544
|||||
247 lYProProAlaLeuGluProProHisAsnLeuProValProGlnLeuThr 263
|||||
5545 GACCTTCGAATCCACAGACAG..... 5564
|||||
264 SerLeuArgProAspIleThrGlnProSerAsnSerThrThrGlyLe 280
|||||
5565 .....AGGCCATCTCTTAACCCCTCTGATTACATACACCTCTGATC 5608
|||||
280 uIleProThrAsnThrProAlaGlnSerSerProGlyValProValLysThrG 297
|||||
5609 ACTCCCACT.....GAGCTTAACATCACTATTAACACAG 5643
|||||
297 lYGlNArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSer 313
|||||
5644 GGGCCAAACTTTTACCTCATCCAGGAGCTTTTCAAGCTCTTACTCTCC 5693
|||||
314 ThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyPr 330
|||||
5694 ACGACTCCAGAGGCTACTCTTCTTGCTTGCTTACTGCTGGGCC 5743
|||||
330 oProThrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysG 347
|||||
5744 ACCTTACTATGAGGAAATGGCTTACAGAGGAGGAATTCATATGACAAAGG 5793
|||||
347 lHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThr 363
|||||
5794 AACATAGAGACCAATGTATCATAGGGGATGCCAAATTAAGCTTACCCTTACT 5843
|||||
364 GluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisGln 380
|||||
5844 GAGGTTTCTGGAAAAAGGACCTGCATAGGAGGCTGCCCATCCACCA 5893
|||||
380 nHisLeuCysThrSerThrValValTyrGluGlnAlaSerGluAsnGlnT 397
|||||
5894 ACACCTTTGTGACCACTGAAGCTTTAAATGCAACCTCTGAGAGTCAAT 5943
|||||
397 yrLeuValProGlyTyrAsnArgTrpPrPalaCysAsnThrGlyLeuThr 413
|||||
5944 ATCTGTATCCGCTTATGACAGGTGTGGGATTAATACTGATTAAC 5993
|||||
414 ProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMe 430
|||||
5994 CTTGTGTTTCCACCTGTTGTTTCAACCAAACTAAAGACTTTTGCCTTAT 6043
|||||
430 tValGlnIleValProArgValTyrTyrHisProGluGluValValLeuA 447
|||||
6044 GGTCCAAATTTGCCCCGGGTGATCTACTATCCGGAAGACATCTCTTG 6093
|||||
447 spGluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeu 463
|||||
6094 ATGATATGATATAGATATATATATGCGCAAAAAGAGAGCCATATCTCCTG 6143
|||||
464 ThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGln 480
|||||
6144 ACACACTGCTGTAATGCTCGATTTGGAGTGGCTGCAAGCGCTGGAGAACAG 6193
|||||
480 yThrAlaIleLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyG 497
|||||
6194 AACGGCTGCTTAATTCACAGAGCCGCAACAGCTGGAGAAAGGACTTAATGA 6243
|||||
497 lLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlnSerVal 513
|||||
6244 ACCTACATGCAATTTGTATAGGAGAGATGCCAACCCCTGAGAAAATATCTGTC 6293
|||||
514 SerAsnLeuGluGlnSerLeuThrSerLeuSerGluValValLeuGlnAs 530
|||||

```

6294 AGTAACCTGGAGGATCCCTAACCTCTTATCTGAAGTGGTCTACAGAA 6343  
 530 nAArgGlyLeuAspLeuPheLeuArgGlyGlyLeuGlyAla 547  
 6344 CAGAAGGGGGTTAGATCTGTATTTTAAAAAAGAGGAGGTTATGTGTAG 6393  
 547 lAlaLeuGlyGlyGlyCysPheTyrValAspHisSerGlyAlaIleArg 563  
 6394 CCTTAAAGAGGAATGCTCTTATGTATGTACATCAGAGGATCATAGA 6443  
 564 AspSerMetAsnLysLeuArgLysLysLeuGlyValArgArgArgGlyAla 580  
 6444 GACCCCATGACAGCCTTAGAGAAAGTTAGAGAGGCGCTGGAAGGGAAG 6493  
 580 gGluAlaAspGlnGlyTyrPheGlyGlyTyrPheAsnArgSerProTyr 597  
 6494 AGAGGCTGACCCAGGGGCTGTGAAGATGGTTCAACAGGCTCTCTTGA 6543  
 597 eTThrThrLeuSerAlaLeuThrGlyProLeuValIleLeuLeu 613  
 6544 TGACACCCCTCTTCTCTCTGACGGGCGCCCTAGTAGTCTCTCTCTG 6593  
 614 LeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheVal 630  
 6594 TTACTTACAGTTGGGCTTGTCTTATTAATAGGTTTGTGCTTGTGTAG 6643  
 630 gGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGln 647  
 6644 AGAAGCAGTGAAGTCAGTCAGATCATGTACTTAGGCACAGATCAAG 6693  
 647 lLeuLeuSerGlnGlyGlyThrAspLeu 656  
 6694 GCCTTCTGAGCCAGGAGAACTGACCTC 6722  
 seq.name: /SDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.AAC67023  
 seq.documentation\_block:  
 ID AAC67023 standard; DNA: 7873 BP.  
 AC AAC67023;  
 DT 27-MAR-2001 (first entry)  
 DE PRV env protein coding sequence SEQ ID NO: 23.  
 KW Xenotransplantation; infectious agent; vaccine; ds.  
 OS Porcine endogenous retrovirus.  
 PN WO200071726-A1.  
 PD 30-NOV-2000.  
 PE 24-MAY-2000; 2000WO-US14296.  
 PR 24-MAY-1999; 99US-0135631.  
 PA (MAYO-) MAYO MEDICAL VENTURES.  
 PI Federspiel MJ;  
 DR WPI: 2001-032041/04.  
 PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 PS Claim 16; Page 112-115; 144p; English.  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from

CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 SQ Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other:  
 alignment\_scores:  
 Quality: 2474.50 Length: 704  
 Ratio: 4.303 Gaps: 12  
 Percent Similarity: 81.676 Percent Identity: 67.330  
 alignment\_block:  
 US-09-171-553b-6 AAC67023  
 Align seg 1/1 to: AAC67023 from: 1 to: 7873  
 1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGlyGlyLys 17  
 5012 ATGCATCCACAGCTTAAGCCGGCCACCTCCGATTCGGGTGGAAAGCC 5061  
 17 oLysArgLeuArgLleProLeuSerPheAlaSerIleAlaTyrPheLeu 34  
 5062 GAAAGACAGAAATCCCTTAAGCTTCCCTTCATCGCGTGTCTTA 5111  
 34 hLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 5112 CTCTGTCATTAACCTCAAGTTAATGTAAGCGCTTGAGACAGCCCG 5161  
 51 AspProHisArgProLeuSerLeuThrPheIleIleAspProAsp 67  
 5162 AACTCCCATAAACCTTATCTCTCAGTGTGTTACTGATCCGCTGAC 5211  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTyr 84  
 5212 AGGTATTAAATATTAACAGACCTCAAGGAGGAGGCTCTGGGAGCTGT 5261  
 84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal 98  
 5262 GGCTGATTAATATATGTCCTTCGATCAGTATCCCTGCTCAATGAC 5311  
 99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115  
 5312 CAGGCCACACCCCGCATGCTGCTGCTAGGGTTTACGTTTACGCC 5361  
 115 oGly...ThrGluLysGluLysTyrCysGlyGlySerGlyGlySerPheC 131  
 5362 AGGACCCCAATATGAAATGAAATATGTTGTGAAATCCTCAGGATTCCTT 5411  
 131 ySarArgArgTyrSerCysValThrSerAsnAspGlyAspTyrPro 147  
 5412 GCAAGCAATGAGCTGGTAACTTCTTAATGATGGAAATGGCA 5461  
 148 lLeuSerLeuGlnAspArgValLysPheSerPheValAsnSer 161  
 5462 GTCTCTCAGCAAGACAGAGTAAGTACTTCTTTGTTAACAAATCCTACG 5511  
 162 .....GlyProGlyLysTyrLys.....MetMet 170  
 5512 TTATATCAATTAATATATGCGCATGAGATGAAAGATGGCAACAGC 5561  
 170 yLeuLeuTyrLysAspLys.....SerCysSerProSerAsp 181  
 5562 GGGTACAAAGATGTAAGCAATTAAGCAATTAAGCTGTCTTTCGTTAGAC 5611  
 182 LeuAspTyrLeuLysIleSerPheThrGlu..ArgLysThrGlyLysTyr 197  
 5612 CTATATTACTTAAATAATAGTTTACAGTGAAGAAACAAACAAATAAT 5661  
 198 SerLysValAspLysTyrTyrGluLeuGlyAsnSerPheLeuLeuTyr 214  
 5662 TCAAAAGTGGTAAATGTAATGTCTTGGGGAATAGTGTAC.....TATG 5705  
 214 lGlyGlyValAla.....GlySerThrLeuThrIleArgLeuArgLle 227



```

|||||..... |||||..... |||||..... |||||..... |||||.....
5706 GAGGCTCTGGGAGAAAGAGATCTGTCTGACTATTCGCCCTCAGAAATA 5755
228 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeu 244
|||||..... |||||..... |||||..... |||||..... |||||.....
5756 GAACATCAGATGAGACCTCGGTTGCTATAGACCAATTAAGGGTTGGC 5805
244 aglGluGlyProProAlaLeuGluProHisAsnLeuProValProG 261
|||||..... |||||..... |||||..... |||||..... |||||.....
5806 CGAACAAGGACCTCCATCCAGACAG..... 5833
261 LIneuThrSerLeuArgProAspLleThrGluProProSerAsnSerThr 277
|||||..... |||||..... |||||..... |||||..... |||||.....
5834 .....AGCCATCTCCTAACCCCTGTGATTACATATACATACACC 5869
278 ThrGlyLeuLeuProThrAsnThrProArgAsnSerProGlyValPro 294
|||||..... |||||..... |||||..... |||||..... |||||.....
5870 TCTGATCAGTCCCACT..... GAGCTTAACATCCTAT 5904
294 LysThrGlyGluArgLeuPheSerLeuIleGlnGlyAlaPheGlnAla 311
|||||..... |||||..... |||||..... |||||..... |||||.....
5905 TAAACAGAGGGCGAAACTTTTAACCTCATCCAGGAGCTTTCAAGCTC 5954
311 LLeuSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 327
|||||..... |||||..... |||||..... |||||..... |||||.....
5955 TTAACCTCAGACTCCAGAGGCTACCTCTTGTGCTTGTCTAGCT 6004
328 SerGlyProProTyrrGlyGluGlyMetAlaLysArgLysPheAsnVal 344
|||||..... |||||..... |||||..... |||||..... |||||.....
6005 TCGGGCCCACTTACTATAGGCAATGGCTAAGAGGAGAAATTCATAGT 6054
344 LThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeu 361
|||||..... |||||..... |||||..... |||||..... |||||.....
6055 GACAAAGAGAACTAGAGACCAATGTACATGGGATGCCAAATTAAGCTTA 6104
361 hLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 377
|||||..... |||||..... |||||..... |||||..... |||||.....
6105 CCTTACTAGGTTCTGGAAAGGACACCTGCATAGGAGTGGTCCCA 6154
378 SerHisGlnHisLeuCysTyrrSerThrValIleGlyGlnAlaSerG 394
|||||..... |||||..... |||||..... |||||..... |||||.....
6155 TCCACACACACCTTTGTACACACACCTGAAGCCTTAATCAGACCTCGA 6204
394 uAsnGlnTyrrLeuValProGlyTyrrAsnArgTrpTrpAlaCysAsnThrG 411
|||||..... |||||..... |||||..... |||||..... |||||.....
6205 GAGTCAGTCTGTGACTGTTATGACAGTGGTGGGATGTATATCTG 6254
411 LLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe 427
|||||..... |||||..... |||||..... |||||..... |||||.....
6255 GATTAAACCCCTGTGTCTCCACCTTGTTCAACCAACCAAAAGACTTT 6304
428 CysValMetValGlnIleValProArgValTyrrTrpHisProGlnGlu 444
|||||..... |||||..... |||||..... |||||..... |||||.....
6305 TCGGTTATGTCCTCAATTTGTCGCCGGGTGTACTACTATCCGAAAAAGC 6354
444 lValLeuAspGlnTyrrAspTyrrArgTyrrAsnArgProLysArgLysPro 461
|||||..... |||||..... |||||..... |||||..... |||||.....
6355 AGTCTGTGATGATATGACTATATATATCGCCAAAGAGAGCCCA 6404
461 aISerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyAl 477
|||||..... |||||..... |||||..... |||||..... |||||.....
6405 TATCCCTGAGCTAGCTGTAAAGCTCGGATGGAGTGGCTCCAGGCTG 6454
478 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlyLysG 494
|||||..... |||||..... |||||..... |||||..... |||||.....
6455 GGAACAGAGAGCGCTGCCCTATATCAGAGCCGACAGCTGAGAAAGG 6504
494 yLeuGlyLeuLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 511
|||||..... |||||..... |||||..... |||||..... |||||.....
6505 ACTTGTACTACATCGAATGTGTAAACGAGATCTCCAAAGCCCTAGAAA 6554
511 lSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 527
|||||..... |||||..... |||||..... |||||..... |||||.....

```

---

```

6555 AATCTGTCAGTAACTGGAGAAATCCCTAACCTCCTTATCTGAAGTGT 6604
528 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlyGlyLe 544
|||||..... |||||..... |||||..... |||||..... |||||.....
6605 CTACAGAACAGAGGGGTTAGATCTGTATTTCTTAAAGAGAGGATTT 6654
544 uCysAlaAlaLeuLysGluGlyCysCysPheTyrrValAspHisSerGly 561
|||||..... |||||..... |||||..... |||||..... |||||.....
6655 ATGTGTAGCTTGAAGAGAGAAATGCTGTTTATATGATCATTTACAGGG 6704
561 lAlleArgAspSerMetAsnLys..... 568
|||||..... |||||..... |||||..... |||||..... |||||.....
6705 CCATCAGAGACTCCATGAACAAGCTTAGAGAAGACTGAGAGGCCCG 6754
568 ..... 568
6755 CGGGTCTGAAACACCAAGAGTGGCTGTGTTCAATTAAGAAGCTGTC 6804
569 .....LeuArgLysLysLeuGluArgArg 577
|||||..... |||||..... |||||..... |||||..... |||||.....
6805 GAAGTAACCGCTGAGCTAAAGCCAGCTTAGAGAAAGTTGGAGAGAGCTC 6854
577 rGArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArg 593
|||||..... |||||..... |||||..... |||||..... |||||.....
6855 GAAGGGAAGAGAAACTACTCAAGAGTGTGAGAGGATGTTCAACAGG 6904
594 SerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuVal 610
|||||..... |||||..... |||||..... |||||..... |||||.....
6905 TCTCCTGTGGTGGCTACCTTCTTCTTCAACAGACCTTAAGT 6954
610 lLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheVal 627
|||||..... |||||..... |||||..... |||||..... |||||.....
6955 CCTCCTCTCTGTACTACAGTGGGCCATGTATATTATTAACAGTTAATG 7004
627 lApheValArgGluArgValSerAlaValGlnIleMetValLeuArgGln 643
|||||..... |||||..... |||||..... |||||..... |||||.....
7005 CCTTCATTAGAGAAAGAACTAGTCAAGTCAGTCAGATCATGTACTTAAGCAA 7054
644 GluTyrrGln 646
|||||..... |||||..... |||||..... |||||..... |||||.....
7055 CAGTAACCAA 7063

```

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT: AAT74811

seq\_documentation\_block:

ID	AAT74811 standard; cDNA; 8060 BP.
XX	AAT74811;
AC	
XX	
XX	
DT	11-FEB-1998 (first entry)
XX	
DE	Porcine retrovirus Tsukuba-1 cDNA.
XX	
KW	Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW	xenotransplantation; infectious; provirus; organ transplant; donor;
KW	activated virus; Tsukuba-1; PCR; ss.
XX	
OS	Porcine retrovirus.
XX	
EH	Location/Qualifiers
FT	86..2002
FT	/tag= a
FT	/note= "putative ENV protein"
FT	3112..4686
FT	/tag= b
FT	/note= "putative GAG protein"
FT	4871..8060
FT	/tag= c
FT	/note= "putative POL protein (partial)"
XX	
PN	W09721836-A1.
XX	
PD	19-JUN-1997.

XX 13-DEC-1996; 96MO-US19680.  
 PF  
 XX 14-DEC-1995; 95US-0572645.  
 PR  
 XX  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PL  
 P1 Fishman JA;  
 XX  
 XX WPI; 1997-332804/30.  
 DR  
 XX  
 XX New nucleic acid from porcine retro:viruses - used for detecting  
 PT viruses in transplant or other tissue and for assessing risk of  
 PT transmitting infection to graft recipient  
 XX  
 PS Claim 1; Fig 1; 128pp; English.

CC This sequence represents the purified porcine retroviral cDNA  
CC sequence of Tsukuba-1 and contains the putative coding regions for viral  
CC proteins Gag, Pol and Env. This sequence and PCR fragments generated  
CC from the sequence (see AAT74812-T74882) could be used to screen organs  
CC for porcine retroviruses prior to xenotransplantation. Transplantation  
CC can increase the likelihood of retroviral activation if intact and  
CC infectious proviruses are present. The porcine retroviral sequence can be  
CC used to generate probes to determine the level (e.g. copy number) of  
CC intact (i.e. potentially replicating) porcine provirus sequences in a  
CC strain of xenograft transplantation donors. It can be used to detect  
CC mutations, genetic lesions or viral recombinants and to determine the  
CC histological localisation of activated retroviruses. Using Polymerase Chain  
CC Reaction DNA Quantitation (PQD) on blood mononuclear cells, infectivity  
CC titration and susceptibility testing can be performed. Ultimately animal  
CC donors without intact porcine retroviral sequences or with a lower copy  
CC number of viral elements could be selected.

Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other

[illegible]

```
alignment_block:
US-09-171-553B-6 x AAT74811 .
```

Align seg 1/1 to: AAT74811 from: 1 to: 8060

[illegible][illegible]

```

408 sasnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 425
1258 TAATAGTGAATTAACCCCTGTGTTCACACCTGGTTTAAACCAACATA 1307
425 ysaspheCysValMetValGlnIleValProArgValIleThrHisPro 441
1308 AAGATTTTGCATTATGTCACAAATGTGTCCCGAGCTGATTAATCTCC 1357
442 GtUGlUValValLeuAspGluTyrAspTyrArgTyrAsnArgProLysAr 458
1358 GAAAAAGCAATCCTTGAATGATATGACTACAGAAATCATCGACAAAAGAG 1407
458 ggIuproValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaV 475
1408 AGAACCAATATCTCGACACTTCTGTGATGCTCGACACTTGAGTGGCAG 1457
475 alGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
1458 CAGGTGTAGAACAGACAGACAGCTGCCCTGTGACGCGGACACAGCAGTA 1507
492 GtUGlUGlyLeuGlyLeuGlnHisAlaIleMetThrGluAspLeuArgAl 508
1508 GAAACAGGACTTATGTAACCTACATGCAATTTGAACAGAAAGATCTCCAAC 1557
508 aleuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerG 525
1558 CCTAGAAAATCTGTCACTACCTGAGAGATCCCTAACCTCTTATCTG 1607
525 lUValValLeuGlnAsnArgGlyLeuAspLeuPheLeuArgGlu 541
1608 AATGTCCTTACAGATAGAGAGGCTTGATTTATTTCTAAAGAA 1657
542 GtUGlUGlyCysAlaIleLeuLysGluGluCysPheTyrValAspHi 558
1658 GGAAGATTATGTGTACCTTGAAGAGAGATCTGTTTATGTGATCA 1707
558 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 575
1708 TTCAGGGGCCATCAGAGACTCATGAACTAGAGAAAGGTTGGAGA 1757
575 rGArgArgArgGluArgGluAlaAspGlnGlyTyrPheGlnGlyTyrPhe 591
1758 AGGTCGAGGAGAAAGAACTACTCAAGGCTGTTTACAGGAGGTTTC 1807
592 AsnArgSerProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLe 608
1808 AACAGGTCTCTGTGTGGCTACCTACTTCTGCTTACAGGACCTT 1857
608 uValValLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgP 625
1858 AATAGTCCTCTCTGTGTACTCAGATGGGCGCATGATTTATTATACAGT 1907
625 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 641
1908 TAATAGCTCTCATTTAGAGAAAGATTAAGTGCAGTCCAGATCATGTACTT 1957
642 ArgGlnGlnTyrGln 646
1958 AGACAAAGATACCAA 1972

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF77725
seq_documentation_block:
ID AAF77725 standard; cDNA: 8060 BP.
XX AAF77725;
AC
XX
XX 23-MAY-2001 (first entry)
XX
XX Tsukuba-1 cDNA.
XX
XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

```

```

XX OS Porcine retrovirus.
XX PN US6190861-B1.
XX PD 20-FEB-2001.
XX PF 13-DEC-1996; 96US-0766528.
XX PR 14-DEC-1995; 95US-0572645.
XX (GENE) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI: 2001-256211/26.
XX DR
XX PT Assessing risk of endogenous retroviruses in clinical practice and in
XX PT xenotransplantation, comprises using probe sequences derived from swine
XX PT or miniature swine retroviral genome
XX PS
XX PS Claim 1; Fig 1; 127pp; English.
XX CC The present invention relates to a method for screening a cell or tissue
XX CC for the presence or expression of a retrovirus (RV), comprising
XX CC contacting a target nucleic acid from the cell or tissue with a second
XX CC nucleic acid from the present invention (e.g. the present sequence or a
XX CC fragment thereof). The method is useful for RV detection and to assess
XX CC graft transplantation risk. Screening of animals allows the elimination
XX CC of donors with active replication of known viruses. Inactive proviruses
XX CC can be detected and inactivated, allowing identification and elimination
XX CC of potential human pathogens derived from swine in a manner not possible
XX CC in the outbred human organ donor population and is important to the
XX CC development of human xenotransplantation.
XX SQ Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;

alignment_scores:
Quality: 2360.00 Length: 656
Ratio: 4.252 Gaps: 11
Percent Similarity: 84.604 Percent Identity: 68.445

alignment_block:
US-09-171-553B-6 x AAF77725
Align seg 1/1 to: AAF77725 from: 1 to: 8060

1 MethisProThrLeuSerArgArgHisLeuProThrArgGlyGlyLupr 17
|||||
86 ATGCATCCACGTTAAGCGGGCGCCACCTCCGATGGGGGTGGAAGGCC 135
17 clyArgLeuArgIleProLeuSerPheAlaSerIleAlaIleThrPheLeu 34
|||||
136 GAAAGACTGAAATCCCTTAACCTTCGCTCCATGCGCTGGTTCCTTA 185
34 hrLeuThrIleThrProGlnAlaSerSerLysArgGlnIleAspSer 50
|||||
186 CTCTGTCAATTAACCTCTCAGACTAATGTATGCCATAGAGAGAGAGCCTG 235
51 AsnProHisArgProLeuSerLeuThrTyrPheLeuIleLeuAspProAspTh 67
|||||
236 AACGCCCATTAACCCCTTATCTCAGCTGTTATTTACTGACGCGGCAC 285
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTyrP 84
|||||
286 AGTATTAAATATCAACAACTCAAGGGGAGGCTCTTTAGGAACCTGGT 335
84 rpproGluLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
|||||
336 GGCCTGATCTATACGTTGGCTCAGATCACTTATCTTACCTGACCTCA 385
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPProGlyTh 117

```

```

386 ...CCCCAGATATCTCCATGCTCAGCAGATTATGTTGCCAGGACC 432
117 rGluYSer...LysTyrCysGlySerGlyLysPheCysAla 133
433 ACCAATATATGAAACATTCGGGAATCCAGAGATTCTTTTGTAAAC 482
133 rGTPSerCysValThrSerAsnAspGlyAspTyrPylsTyrProIleSer 149
483 AATGAGACTGTACCTCTATGATGATATGTAATGGCAACCTCT 532
150 LeuGlnAspArgValLysPheSerPheValAsn...SerGlyProGln 164
533 CAGCAGAGATAGGGTAAGTTTCTTATGTGCAACACCTATACCAAGCTCTGG 582
164 LysTyrLysMetMetLysLysLysLysLysLysSer...CysSerP 179
583 ACAATTTATATTCCTGACCTGATTAAGAACGGAAGCCCAAGTCTCTC 632
179 rOSerAspLeuAspTyrLeuLysLleSerPheThr...GluArgLysThrG 195
633 CTTCAGACCTGATTAACCTAAATAAGTTTCACCTGAGAAAGAAACAA 682
195 LysTyrSerLysValAspLysTyrPylsTyrGluLeuGlyAsnSer...PheLe 211
683 GAAATATCTCTAAATGGGTAATGTAATGCTTGGGAATGTAAT.. 730
211 uGluTyrGlyGlyGlyAla...GlySerThrLeuThrIleArgL 225
731 ...TATGAGGCTCGGGTAACACACAGGCTCCATCTTAACCTATTCGCC 776
225 euArgIleGluThrGlyThrGluProValAlaMetGlyProAspLys 241
777 TCMAATATTAACCG...CTGAGCCTCCAAATGGCTATGAGCAAAATGCG 823
242 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 258
824 GTCTTGACGGCTCAAGACCCCAACCCAGAGACCA..... 859
258 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSerA 275
860 .....GGACCATGCTCTA 872
275 snSerThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291
873 ACATACCTCTGATCAGACCCCACTGATCTAGCAGCAGCACT..... 916
292 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPh 308
917 .....AATATGGGGGCAAACTTTTATGCTCATCCAGGAGCTTT 957
308 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTyrLeu 325
958 TCACGCTCTTACTCCACGACTCCAGAGCTACCTCTTGTGTGGCTAT 1007
325 yLeuSerSerGlyProProTyrTyrGluGlyMetAlaLysGluArgLys 341
1008 GCTTAGCTTGGGCCCTTACTATGAAAGGCTAGAGAGGAGGAAA 1057
342 PheAsnValThrLysGluHisArgAsnGlnCysThrTyrPylsSerArgAs 358
1058 TTCATATGTGACAAAGACATAGACCAATGSCATGSGGATCCCAAAA 1107
358 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 375
1108 TAAAGCTTACCTTACTGAGGTTTCGGAAGGCACTGCATAGAGAAAG 1157
375 laProProSerHisGlnHisLeuGlySerTyrSerThrValValTyrGln 391
1158 TTCCCCCATCCACCAACACCTTTGTAAACCACTGAAGCCCTTAATATA 1207
392 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTyrPylsAlaCy 408
.....
1208 ACCTGTGAAGTCAATATCTGTACTGCTTATGACAGGTGGTGGCCATG 1257
408 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 425
1258 TAATATGAGATTAACCCCTTGTGTTCACCTGTGGTTTAAACCAATA 1307
425 yAspPheCysValMetValGlnIleValProArgValTyrTyrHisPro 441
1308 AAGATTTTTCATATGCTCAATATGTTCCAGAGTATTAATCTATCC 1357
442 GluGluValValLeuAspGluTyrAspTyrArgTyrAsnArgProLysAr 458
1358 GAAACACCAATCTCTGATGATATGACTACGAAATATCATGACAAAAAG 1407
458 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAla 475
1408 AGAACCATATATCTGTACACTGTGTGATGCTGCGACTGGAGTGGCAG 1457
475 aLgLYValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
1458 CAGGTGTAGAGACAGCAAGACGCTGCCGTGACAGGAGCCACAGCACTA 1507
492 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 508
1508 GAAACAGAGACTTACTATACCTACATCGAATGTGTAACAGAAATCTCCA 1557
508 aLeuLysGluSerValSerAsnLeuGluGlnLysLeuThrSerLeuSer 525
1558 CCTAGAAATATCTGTCTAGTAACCTGAGAGAAATCCCTTAACCTCTTAT 1607
525 luValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 541
1608 AAGTAGTCTTACAGAAATGAAGAGGTTAGATTATTAATTTCTPAAGAA 1657
542 GlYLYLeuCysAlaAlaLeuLysGluGluCysCysPheTyrValAspH 558
1658 GGAGGATTAATGCTGTAGCTTGAAGAGAGATGCTGTTTATATGATCA 1707
1708 TTCAGGGGCGCATCAGAGACTCCATGAAACAACTTAGAGAAAGCTTGGAG 1757
575 rGArgArgArgGluArgGluAlaAspGlnGlyTyrPylsGlyTyrPyls 591
1758 AGCCTCAAGAGGAAAGAACTACTCAAGGAGGTTTGAGGAGTGGTTC 1807
592 AsnArgSerProTyrPylsThrLeuLeuSerAlaLeuThrGlyProLe 608
1808 AACGAGTCTCTGTGTGGCTACCTACTTTCGCTTAAACAGACCTT 1857
608 uValValLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgP 625
1858 AATGTCCTCTCTCTGTACTACAGTGGGCAATGATATTAATTAACAAAG 1907
625 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 641
1908 TAATGCTTCATTTAGAAACGAATTAAGTCACTCCAGATCATGTRACT 1957
642 ArgGlnGlnTyrGln 646
1958 AGACAAACAGTACCAA 1972
seq_name: /sids2/gcgdata/geneseq/geneseqn/NA1997.DAT:AA174884
seq_documentation_block:
ID AA174884 standard; cDNA; 7892 BP.
XX
AC AA174884;
XX
DT 09-FEB-1998 (first entry)
XX
DE Miniature swine retrovirus cDNA.
XX

```

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
 KW xenotransplantation; infectious; provirus; organ transplant; donor;  
 XX activated virus; PCR; ss.  
 OS Porcine retrovirus.  
 XX  
 FH Location/Qualifiers  
 FT CDS 585..2159  
 FT /tag= a  
 FT mat\_peptide 585..2156  
 FT /tag= b  
 FT /note= "putative GAG protein"  
 FT CDS 2307..5741  
 FT /tag= c  
 FT mat\_peptide 2307..5741  
 FT /tag= d  
 FT /note= "putative POL protein"  
 FT CDS 5620..7536  
 FT /tag= e  
 FT mat\_peptide 5620..7533  
 FT /tag= f  
 FT /note= "putative ENV protein"  
 XX  
 PN MO9721836-A1.  
 PD 19-JUN-1997.  
 XX  
 PF 13-DEC-1996; 96MO-US19680.  
 XX  
 PR 14-DEC-1995; 95US-0572645.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Fishman JA;  
 XX  
 DR WPI; 1997-332804/30.  
 DR P-PDB; AAW32096-W32098.  
 XX  
 PT New nucleic acid from porcine retroviruses - used for detecting  
 PT viruses in transplant or other tissue and for assessing risk of  
 PT transmitting infection to graft recipient  
 XX  
 PS Claim 22; Fig 3; 128bp; English.  
 XX  
 CC This cDNA sequence represents a porcine retrovirus from miniature swine  
 CC containing the putative coding regions for viral GAG, POL and ENV  
 CC proteins. This sequence and PCR fragments generated from it  
 CC (see AAT74812-174882) can be used to screen organs for the presence of  
 CC porcine retroviruses prior to xenotransplantation. Transplantation can  
 CC increase the likelihood of retroviral activation if intact and  
 CC infectious proviruses are present. The porcine retroviral sequence can be  
 CC used to generate probes to determine the level (e.g. copy number) of  
 CC intact (i.e. potentially replicating) porcine provirus sequences in a  
 CC strain of xenograft transplantation donors. It can be used to detect  
 CC mutations, genetic lesions or viral recombinants and also to determine  
 CC the histological localisation of activated retrovirus. Using Polymerase  
 CC Chain Reaction DNA Quantitation (PQ) on blood mononuclear cells,  
 CC infectivity titration and susceptibility testing can be performed.  
 CC ultimately animal donors without intact porcine retroviral sequences or a  
 CC lower copy number of viral elements could be selected.  
 XX  
 SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

alignment\_scores:  
 Quality: 2353.00 Length: 656  
 Ratio: 4.232 Gaps: 12  
 Percent Similarity: 84.756 Percent Identity: 68.598

alignment\_block:  
 US-09-171-553B-6 x AAT74884 ..  
 Align seg 1/1 to: AAT74884 from: 1 to: 7892

1 MethisProthLeuSerArgArgHisLeuProthArgGlyGlyLeu 17  
 5620 ATGCATCCCAAGTTAAACCGCGCCACCTCCGATTCGGGTGGAAGCC 5669  
 17 olysArgLeuArgIleProLeuSerPheAlaSerIleAlaIlePheLeu 34  
 5670 GAAAGACGTGAATCCCTTAAGCTTCCTCCATCGCGGTGTCCTTA 5719  
 34 hrieuthrIleThrProGlnAlaSerSerIleArgLeuLeaspSerSer 50  
 5720 CTCTGTCAATTAACCTCTCAGACTAATGATGCGCATAGGACAGCCCTG 5769  
 51 AsnProHisArgProLeuSerLeuThrPheLeuIleLeaspProAsp 67  
 5770 AACTCCCATTAACCCCTTATCTCTCAGCTGTAATTACAGATCCGGCAC 5819  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrPT 84  
 5820 AGCTATTAAATACACACACTCAAGGGAGGCTCTCTTAGAAGACCTGCT 5869  
 84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 5870 GGCTGATCTATACGTTTGCCTGAGATCAGTATTCTTAGTCTGACCTTA 5919  
 101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGly 117  
 5920 ..CCCCAGATATCCTCATGCTCAGGATTTTATGTTGGCCAGGACC 5966  
 117 rGlyLysGlu...LysTrpCysGlyGlySerGlyGlySerPheCysArg 133  
 5967 ACCAAATTAATGAAACATGGGGAATCCAGAAATCTTTCTTTGTAAC 6016  
 133 rGTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPheSer 149  
 6017 AATGGAAGTGTGTAACCTCTAATGATGATATGGAAATGGCAACCTT 6066  
 150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProG 164  
 6067 CAGCAGAGATAGGGTAGTATTCTTATATGCAACACTAATACAGAGCTG 6116  
 164 yLysTrpLysMetMetLysLeuTrpLysAspLysSer.....CysSerP 179  
 6117 ACAATTAAATTACCTGACCTGATTTAGAACTGGAAGCTGCTCTTC 6166  
 179 rSerAspLeuAspTrpLeuLysIleSerPheThr..GluArgSerThg 195  
 6167 CTTCAGACCTAGATTACCTTAAATAATAGTTTCTACTAGAAAGAAACA 6216  
 195 yLysTrpSerLysValAspLysTrpTrpGlyLeuGlyAsnSer..Phe 211  
 6217 GAAATATCTCTTAATAATGGGTAAATGTATGCTGTGGGGAATGGTAT 6264  
 211 uLeuTrpGlyGlyGlyAla.....GlySerThrIleThrIleArgL 225  
 6265 ....TATGAGGCTCGGTGTAACCAACAGGCTCATTAATATTCGCC 6310  
 225 euArgIleGlyThrGlyThrGlyThrProProValAlaMetLysProAsp 241  
 6311 TCATAATTAACCG...CTGGAACCTCCATAGGCTATAGGACCAATATAC 6357  
 242 ValLeuAlaGluGlnGlyProProAlaLeuGluProHisAsnLeuPr 258  
 6358 GTCTTGACGGGTCAAGAGACCCCAACCAAGGACCA..... 6393  
 258 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSer 275  
 6394 .....GAGCCATCTCTTA 6406  
 275 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291  
 6407 ACATTAACCTTGATCAGACCCCACTGAGTCT.....AACAGC..... 6444

292 ValProValIysThrGlyAlaLeuPheSerLeuIleGlnGlyAlaPh 308  
6445 ...ACGCCTAAATGGGGCAAAACTTTTACCGTCATCCAGGACCTT 6491

308 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuc 325  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6492 TCAGCTCTTAACCTCCACGACCTCCAGAGGCTACCTCTTGTTGGCAT 6541

325 ylsLeuSerSerGlyProProTyrrTyrrGluGlyMetAlaLysLuarGlys 341  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6542 GCTTAGCTTGCGGCCACCCTTACTTAAGAAGGAATGGCTAAGAAGGAAAA 6591

342 PheAsnValThrIysGlnHisArgAsnGlnCysThrTrpGlySerArgAs 358  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6592 TTCATATTGCAAAAGACATAGACCAATGACCATGTGGGGATGCCAAAA 6641

358 nLysLeuThrLeuThrGluValSerLylsGlyThrCysIleGlyLysA 375  
6642 TAAGCTTACCCTTACAGAGGTTCTGTGAAAAGGACACCTGCATAGGAAGA 6691

375 lApProSerHisGlnHisLeuCysTyrSerThrValAlaTryGluGln 391  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6692 TTCCCCCATCCCACAACACCTTTGAACACACAGTGAAGCCCTTTAATCAA 6741

392 AlAspSerGlnAsnGlnIlyrLeuValProGlyTyrAsnArgTrpTrpAlaCys 408  
6742 ACCCTGAGAGTCAAAATATGTGTACTGTGTATGACAGGTGGTGGGCATG 6791

408 sAsnThrGlyLeuThrProCysValSerThrSerAlaPheAsnGlnSerL 425  
6792 TAACTACTGATTAACCCCTTGtGTTCCTCACCTGGTGTTTTAACCAACTA 6841

425 ysAspPheCysValMetValGlnIleValProArgValTyrTrpHisPro 441  
6842 AAGATTTTTGCATTATGATGTCACAAATGTCTCCCGAGTGATTAATACTCC 6891

442 GluGlnValValLeuAspGlnIlyrAspTyrArgTyrAsnArgProLysArg 458  
6892 GAAAAAGCAATCCCTGATATAATATGATCAACAATAATCATCGACAAAGAAC 6941

458 gGIUpProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaV 475  
6942 AGAACCCATATCTCTACACCTTGCTGTATGCTCGGACTTGGAGTGGCAG 6991

475 alGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491  
6992 CAGGTGTAGGACAGAGAACAGCTGCCCTGTGTACAGGACCAAGAGCTA 7041

492 GluLysGlyLeuGlyLysLeuHisAlaAlaMetThrGluAspLeuArgAl 508  
7042 GAACACGAGCTATGTAATCTTCACTGCAATGTACACAGAAATCTCCAAGC 7091

508 alLeuLysGlnSerValSerAsnLeuGlnGlnSerLeuThrSerLeuSerG 525  
7092 CCTAGAAAAATCTGTACATTAACCTGGAGGAATCCTTAACCTTCATTATG 7141

525 lValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgLysL 541  
7142 AAGTAATCTCTACAGATTAAGAAGGTTAAGATTATTTCTTAAGAAAGAA 7191

542 GlyGlyLeuCysAlaAlaAlaLeuLysGlnGlyCysCysPheTyrValAspH 558  
7192 GGGAGGATTATGTGTACCTTGAAGAGAGGAAATCTTTTATGTGGATCA 7241

558 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGlnU 575  
7242 TTCAGGGCCATCAGAGACTCATATACACAGCTTAGAAGAAAGGTGGAGGA 7291

575 rGrArgArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPhe 591  
7292 AGCGTGAAGGAGAAAGAACTACTCAAGGGTGGTTTGAAGGATGGTTC 7341

592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 608

```

US-09-171-553B-6 x AAF77727 ...
Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 MethisProthLeuSerArgArgHisLeuProthArgLysGlyLysIupr 17
|||||.....
5620 ATGCATCCACAGTAAACGGCGCCACCTCCCGATCGGGGTGGAAAGCC 5669

17 OlyArgLeuArgLLeuProLeuSerPheAlaSerIleAlaIlePheLeu 34
|||||.....
5670 GAAAGACTGAAATCCCTTAACCTTCGCTCCATCGCGGTGGTCTCTTA 5719

34 hLeuThrLLeuThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||.....
5720 CTGTGCAATACCTCTGACACTATGATGATGCCATAGAGACAGGCTG 5769

51 AsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspTh 67
|||||.....
5770 AACTCCCATTAACCTTATCTCTCAGCTGGTTATCTAGCTCCGCGCAC 5819

67 rGlyValThValAsnSerThrArgGlyValAlaIleProArgGlyThrTrp 84
|||||.....
5820 AGGTATTAAATATCAACAACTCAAGGAGGCTCTTTAGGAACCTGGT 5869

84 rProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||.....
5870 GGCTGATCTATACGTTTGGCTCAGATCAGTTATCTCTAGTCCAGCTCA 5919

101 ThrProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyTh 117
|||||.....
5920 ...CCCCAGATATCCTCCATGCTACGAGATTATGTTTGGCCAGAGCC 5966

117 rGlnLysGlu...LysTrpCysGlyGlySerGlyGluSerPheGlyArg 133
|||||.....
5967 ACCAATATATGAAAAACATTCGGGAATCCAGAGATTCTTTGTATAAC 6016

133 rGTPSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysSer 149
|||||.....
6017 AATGGAACCTGTAACTCTTAATGATGATATTTGMAATGGCCAACTCT 6066

150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProG 164
|||||.....
6067 CAGCAGATAGGTAAAGTTTCTTATGTCACACCTATACCACTCTGG 6116

164 yLysTrpIleMetLeuLysLeuTrpLysAspLysSer.....CysSerP 179
|||||.....
6117 ACAATTTAATTAACCTGACCTGATTAGAACTGAAAGCCCAAGGCTCTC 6166

179 rSerAspLeuAspTrpLeuLysIleSerPheThr..GluArgLysTrpG 195
|||||.....
6167 CTTCAGACCTAGATTACTTAAATAATAGTTTCACTGACAAAGAAACAA 6216

195 lYlYsTrpSerLysValAspLysTrpTrpGluLeuGlyAsnSer..Phe 211
|||||.....
6217 GAAATATCTCTAAATAATGGTAAATGATGTTCTTGGGGAATGATATAT 6264

211 uLeuTrpLysGlyGlyAla.....GlySerThrLeuThrIleArgL 225
|||||.....
6265 ...TATGGAGGCTCGGGTAAACAACGAGCTCATCTTCTACTATCTGCC 6310

225 euArgLLeuGluThrGlyThrGluProProValAlaMetGlyProAspLys 241
|||||.....
6311 TCAAAATATAACAG...CTGGAGCTTCAATGGCTATATAGACCAATATG 6357

242 ValLeuAlaGluGlnGlyProProAlaLeuGluIleProHisAsnLeuP 258
|||||.....
6358 GTCTGAGCGGTCAAAAGACCCCAACCAAGACCA..... 6393

258 oValProGlnLeuThrSerLeuArgProAspLLeThrGlnProProser 275
|||||.....
6394 .....GGACCATCTCTTA 6406

275 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291

|||||.....
6407 ACATTAACCTCTGATATCAGACCCACAGTACGCT.....AACAGC..... 6444

292 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaP 308
|||||.....
6445 ...ACACATATAATGGGGCAAAACCTTTTATGCTCATCTACAGGACCTTT 6491

308 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeu 325
|||||.....
6492 TCAGGCTCTTAACCTCCAGACTCCAGAGGCTTACTCTTTTGTGGCTAT 6541

325 yLysLeuSerSerGlyProProTrpTrpGlyGlyMetAlaLysGluArgLys 341
|||||.....
6542 GCTTACTCTGGGCGCCACCTTACTATGAAAGATGGCTGAAAGAGGANA 6591

342 PheAsnValThrLysGlnHisArgAsnGlnCysThrTrpGlySerArgAs 358
|||||.....
6592 TTCAATGTGCAAAAGAACATAGAGACCAATGACATGGGGAATCCCAAA 6641

358 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 375
|||||.....
6642 TAAGCTTACCTTACTGAGGTTTCTGGAAGGACCTGCATAGAAAG 6691

375 lAbProProSerHisGlnHisLeuCysTrpSerThrValValTrpGluGln 391
|||||.....
6692 TTCCCCCATCCCAACAAACCTTTGTAAACCACTGAAGCCTTATATCA 6741

392 AlAspGlnAsnGlnTrpLeuValProGlyTrpAsnArgTrpPrpLacy 408
|||||.....
6742 ACCTGTAGAGTCAATATCTGATCTGTTATGACAGGTTGGGGCATG 6791

408 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 425
|||||.....
6792 TAATACTGATTAACCCCTTGTGTTTCCACCTGTTTAAACCAACATA 6841

425 yAspPheCysValMetValGlnIleValProArgValTrpTrpHisPro 441
|||||.....
6842 AAGATTTTTCATATATGTCACAAATGTTCCCGAGATGATATCTATCC 6891

442 GluGluValValLeuAspGluTrpAspTrpArgTrpAsnArgProLysAr 458
|||||.....
6892 GAAAGAACCAATCCTTGATGATGATGACTACAGAAATCAGCAACAAAG 6941

458 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThAlaV 475
|||||.....
6942 AGAACCATATCTCTGACACTTCTGTGATGCTCGGACCTTGGAGTGGCAG 6991

475 aLgLYValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
|||||.....
6992 CAGGTGAGAACAGACAGAACGCTGCCCTGGTCAAGGAGCACAGCAGTA 7041

492 GluLysGlyLeuGlyGluLeuHisAlaIleMetThrGluAspLeuArgAl 508
|||||.....
7042 GAAACAGAGCTTGTATTAACCTACATGCAATTTGTACAGAGATCTCCAAG 7091

508 aLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerG 525
|||||.....
7092 CCTAGAAAATCTGTCTGATCTGATCTGAGGAGATCCCTAACCTCTTATCTG 7141

525 lUValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlu 541
|||||.....
7142 AAGTAGCTCTACAGAAATAGAGAGGTTAGATTATTAATTCYAAAAGAA 7191

542 GlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheThrValAspH 558
|||||.....
7192 GAGGATATATGTATGCTTGAAGAGAGATGCTGTTTATATGTGATCA 7241

558 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 575
|||||.....
7242 TTTCAGGGGCAATCAGAGACTCCATGAACAAGCTTAAAGAAAGTTGGAG 7291

575 rGArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 591
|||||.....

```

7292 AGCGTGAAGGAAAAAGAATCACTCACAAGGGTGTTGAGAGATTGTTC 7341

592 AAsArSSePProTrIpMeThrThrLeuSeusSerAlalaLeuthrglyProle 608  
||||| | :||:::||:  
7342 AAAGGtCrCTTTgTgGTtgGCcAcCCtACTTCCTgcTTTTAAACGACCCTT 7391

608 uValVaLlAuLeuLeuLeuLeuThVaiGlYPrOcYSLeuIleaSnRgp 625  
::: ||::||  
7392 AATAgrtcctcgccgtttactcacagttgggccatgatatttatTAACAAGT 7441

625 heVALalabhevalargcluaravalseralavalglntlemetvallen 641  
::::::::::|||::|  
7442 TAttTccCTTCAtTAGAgAACgaATMAgTGcAGTCCAATCATGTGActT 7491

642 ArgGLInLTyrGLn 646  
||| |  
7492 AGCACACACTACCaa 7506

seq\_name: /SID52/gcndata/geneseq/geneseqn/NA2000.DAT.AA251087

seq\_documentation\_block:

ID AA251087 standard; DNA; 2000 BP.

xx AC AA251087;

xx DT 05-JUN-2000 (first entry)

xx DE Contiguous DNA of porcine endogenous retrovirus-D (PERV-D) env region.

xx KW Porcine endogenous retrovirus; PERV-D; virucide; prevention; vaccine;  
KW pIs; diagnosis; infection; xenotransplantation; antibody; env region; ds  
Sus scrofa.

OS Sus scrofa.

NN WO200011187-A1.  
NN PD 02-MAR-2000.  
XX PF 18-AUG-1999; 99WO-US19053.  
XX PR 18-AUG-1998; 98US-0097015.  
PA (BIOT-) BIO TRANSPLANT INC.  
PI Banerjee PT, Patience C, Andersson GK;  
DR WPl: 2000-224704/19.  
XX XX

PT Porcine retroviral PERV-D polypeptides for diagnosing porcine  
PT retroviral infections in humans after xenotransplantation -  
PS PS  
XX Example 4; Fig 11; 119pp; English.

The present DNA is the contiguous sequence of porcine endogenous  
retrovirus-D (PERV-D) env region. It is compiled from the DNA fragments  
obtained from the 5' end, extended proline rich region and 3' end of  
CC PERV-D env region. It is isolated from the porcine kidney, PK15 cell line  
CC (ATCC No.CCL-33). PERV-D DNA has 79% homology to a portion of PERV-C and  
has virucidal activity. PERV-D sequence is useful for prevention or  
diagnosis of infection of human tissues by porcine retroviruses after  
xenotransplantation procedures. PERV-D polypeptide may be used to produce  
specific antibodies, that can be administered as vaccines to create  
passive immunity. The DNA can be used as a hybridisation probe or primer  
for isolation purposes.

Sequence 2000 BP; 667 A; 468 G; 366 C; 499 T; 0 other:

[illegible]

```

alignment_block:
US-09-171-553B-6 x AA251087 ..
Align seg 1/1 to: AA251087 from: 1 to: 2000

```

1 MethionineProteinLeucineSerineArginineHistidineProthionineArginylglycylleu  
43 AFNACATCCACGTTAACCGCGCCGACCTCTGATTCAGAGGTCAAAAGACC 92  
17 OlysatrigleuArlgleProleuSerPhealaSerllealatrpheleut 34  
34 hleuthThleThProhlaIaSerSerlyArgleuIleasrPser 50  
143 CTCCTGCAATACCTCTCAACACTAATGGATGACATAGCAACAGCCTG 19  
51 AsnProhlaIaProleuSerleuthThrpheulleIleasProAspTh 67  
193 AACTGCCATAAACCTTATCTGACCTGGTTAATTAACGACTCGAAC 24  
67 RglyAlthValaIasnSerThArgglyValaIaProArlglyThTrp 84  
243 AGGATTAATATATCCACAGCCCTCGAGGGAGGCTCTTACAAACCTGGT 29  
84 rProGluIleHlaIspheCysLeuArlgleuIleAsnProAlaValysSer 100  
293 GGCTGATCTATATGCTGCTTACGATCAGTCAATTCCTAGTGTGACCTCA 34  
101 ThProhlaIaIaValaIaSerThArgSerThArglyPheThCysCysProGlyTh 117  
343 ACC...CCAGATATCCCTCGCTGCTTACGAGATTTATGTTGGCCAGGACC 38  
117 RgluIlyGlu...LysThrcysGlyIlyserIlyleuSerPheCysArg 133  
390 ACCAAATAATGAGAAACACATGAGAAATCTAGAAATTTCTTTTAAAC 43  
133 rGrPserCysValThThSerAsnAspGlyAspTrpIlyTrpProIleSer 149  
440 AATGACACTGTGTAACCTCTAATGATGAGAAATCGAAATGGCCAACTCT 48  
150 LeuGlnAspArgValIlyspheSerPheValaIaSerGlyProGlyIlysty 166  
490 CTGAGAGATAGGGTAGACTTTTCTTATGTCACAC...CCCAATTAACA 53  
166 RlyMetMetIlyleu...TyrIlyAspIlySer...CysSerProSer 181  
534 CCGGACTCGAAACACTCATACAGGACCATTCCTGGGTGTTTCCCTCAG 56  
181 spleuAspTrpIleuIlyIleSerPheThrglu...ArgIlyThrglyIlyST 197  
584 ACCTAGATATATCTTAATAATAGTTTCACGCAAAAAAACAAGAAAT 63  
197 yIserIlyValaIlyAspIlySTrPTrpIlyGluIleuGlyIasnSerPheleuLeuTyr 213  
634 ATCTCAAAAAGATTAATGTTATGCTCTCGGGAAATAAAT...ATATTTATAC 68  
214 .GlyIlyGly...AlaGlySerThleuthThlleIleIlyleuArlgle 228  
681 AGGTGGGACAGCAACACAGGCTCATATTTAACCATCCGACTTAATAATA 79  
228 IuthrglyThrgluProProValaIaMetGlyProAspIlyValleuAla 244  
730 ...AGCGAGTAGAGCTCCCAATAGGCTAATAGGACGAATACGGTCTTAACG 77  
245 GluGlnGlyProProGluIleuGluProProhlaIasnleuProValProG 261  
778 GATCAAGAAAGACCCCAACCCAGACACATCTCTGATATA..... 816  
261 nleuthTrleuArlProAspIleThGlnProProSerAsnSerThTrp 278  
817 .....ACTCTTAATTAAGACCCACGCTAG...CTTAACGACGACGA 853



```

278 hrglyleu1leuprothrasnthrProArgAsnSerProGlyValProVal 294
||
854 CT..... 855
295 LysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 311
||||||| ..... 311
856 AAAACGGGACAAACCTTTTATGCTCATCCAGGAGCTTTCAAGCTCC 905
311 eAsnSerThrAspProAspAlaThrSerSerCysThrLeuCysLeuSerS 328
||||||| ..... 328
906 TAACCTCAGAGCTCCAGAGGCTACCTCTTGTGGCTTGCTTAACCTT 955
328 eGlyProProTyrTyrGlnGlyMetAlaLysGlnArgLysPheAsnVal 344
||||||| ..... 344
956 CGGGCCACCTTACTATAGAAATAAGGCTAAAGAAATAATTCATGTG 1005
345 ThrLysGlnHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuTh 361
||||||| ..... 361
1006 ACMAAAACATAGAGACCAATGTACATGGGATCCCAAAATACGCTTAC 1055
361 rLeuThrGlnValSerGlyLysGlyThrCysIleGlyLysAlaProProS 378
||||||| ..... 378
1056 CCTTACTGAGGTTTCTGGAAAGACACCTGCATAAAAGGTTCCCAT 1105
378 eHisGlnHisLeuCysTyrSerThrValValTyrGlnGlnAlaSerGln 394
||||||| ..... 394
1106 CCCACCAACACCTTACAAACCACTGAAGCCTTAAATCAAAACCTCTGAG 1155
395 AsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGln 411
||||||| ..... 411
1156 AGTCATATCTGTGTAAGCTGTATGACAGAGTGGTGGCATGTAAATACGTG 1205
411 yLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheC 428
||||||| ..... 428
1206 ATTAACCCCTGTGTGTTCCACCTGTGTTTCACCAAACTAAAGCTTTT 1255
428 yValMetValGlnIleValProArgValTyrTyrHisProGlnGlnVal 444
||||||| ..... 444
1256 ACATTTATGTCCAAATGTGCCCGAGTATATTACTATCCCAAGAAACA 1305
445 ValLeuAspGlnTyrAspTyrArgTyrAsnArgProLysArgGlnProVa 461
||||||| ..... 461
1306 ATTCGTGATGATATGATTACAGAACCATCGACAAAGAAACCAAT 1355
461 lSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyValG 478
||||||| ..... 478
1356 ATCCCTGACACTCGCAGTAATGCTCGAGCTCGAGTGTATACAGGTGTA 1405
478 lYThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlnLysGly 494
||||||| ..... 494
1406 GAACAGGAACCTGACCTTATGTTACAGGACCTCAGACGCTAGAAACAGA 1455
495 LeuGlyGlnLeuHisAlaAlaMetThrGlnLysPheuArgAlaLeuLysGln 511
||||||| ..... 511
1456 CTTAGTACATCATCAATTTGTACAGAAATCTCCAAAGCCCTAAAGAAA 1505
511 uSerValSerAsnLeuGlnGlnSerLeuThrSerLeuSerGlnValValL 528
||||||| ..... 528
1506 ATCTGTGAGTAACTCGAAATAATCCCTAACCTCTATCTGAGAGTTC 1555
528 euGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGlnGlyGlyLeu 544
||||||| ..... 544
1556 TACAGAAATAAAAGGTTAGATTATTTCTTAAATAAAAGAAAGATTA 1605
545 CysAlaAlaLeuLysGlnGlyCysCysPheTyrValAspHisSerGlyAl 561
||||||| ..... 561
1606 TGTGTAGCCTTAAAGAGAAATGCTGTTTATGTAGATCATTCACAGGGGC 1655
561 aIleArgAspSerMetAsnLysLeuArgLysLysLeuGlnArgArgArg 578
||||||| ..... 578
1656 CATCAGAGACTTCATGACAAAGCTTAAATAAAAGGTTGAGAAACGTCGAA 1705
578 rGlnArgGlnAlaAspGlnGlyTyrPheGlnGlyTyrPheAsnArgSe 594

```

```

||||||| ..... ||| ||||| ..... |||||
1706 GGGAAAGGAAACCTTACTCAAGATGTTTAAAGATGTTCAACAGGTC 1755
594 rProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValL 611
||||||| ..... 611
1756 TCCTTGGTTGGCTACCTTCTTACTTTAACAGGACCTTAAATAGTCC 1805
611 euLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 627
||||||| ..... 627
1806 TCCTCTGTTACTCAGAGTTGGCCATGTATTATTAACAGTTAATGCC 1855
628 PheValArgGlnArgValSerAlaValGlnIleMetValLeuArgGlnGln 644
||||||| ..... 644
1856 TTCATTGAAACGCAATAGTGCAGTCAATCATGTGTTAGACAAACA 1905
644 nTyrGln 646
|||||||
1906 GTACCAA 1912

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 03:02:54 ; Search time 624.99 Seconds

(without alignments)  
4054.870 Million cell updates/sec

Title: US-09-171-553b-9

Perfect score: 2956  
Sequence: 1 Tgcttttaagggttaggaac.....aaaaaaaaaaaaaaaaaa 2956

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 segs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2896.8	98.0	8196	AAV09699	Porcine retrovirus
2	2896.4	98.0	8209	AAV09700	Porcine retrovirus
3	2889.4	97.7	3482	AAV82749	Pig endogenous ret
4	1970.8	66.7	1974	AAV09703	Porcine retrovirus
5	1882.8	63.7	7333	AAV77726	Defective retrovir
6	1809.6	61.2	7393	AAV74883	Porcine retrovirus
7	1563.4	52.9	2462	AAV82748	Pig endogenous ret
8	1552.2	50.4	4918	AAV67022	PERV env protein c
9	1489.2	48.5	8132	AAV77727	Nucleotide sequenc
10	1433.8	45.6	7873	AAV67023	PERV env protein c
11	1347.8		7892	AAV74884	Miniature swine re

12	1231.8	41.7	8060	AAV74811	Porcine retrovirus
13	1231.8	41.7	8060	AAV77725	Tsukuba-1 cDNA. P
14	1009.8	34.2	1980	AAV67018	PERV-1-15 env prot
15	973.8	32.9	2000	AAV51087	Continuous DNA of
16	719.4	24.3	7362	AAV67019	PERV env protein c
17	710.4	24.0	1493	AAV51086	3'DNA of porcine c
18	622.8	21.1	4402	AAV67020	PERV env protein c
19	600.2	20.3	3320	AAV09698	Porcine retrovirus
20	455.2	15.4	6312	AAV09691	Plasmid pBdelPASA
21	451.2	15.3	8088	AAV65540	Complete nucleotid
22	451.2	15.3	8535	AAV73731	GALV SEAVO genome.
23	446.6	15.1	2129	AAV45343	DNA encoding the e
24	433.8	14.7	2055	AAV89076	GALV envelope glyco
25	430.4	14.5	2129	AAV45344	DNA encoding the e
26	419.4	14.2	1998	AAV89080	GALV envelope glyco
27	419.4	14.2	2004	AAV65507	DNA sequence of a
28	372.2	12.6	8655	AAV65750	Mus dunali endogeno
29	357.8	12.1	10970	AAV73735	Retro virus vector
30	343.4	11.6	698	AAV51073	5'DNA of porcine e
31	313.4	10.6	1941	AAV83504	Amphotrophic MLV r
32	309.2	10.5	1923	AAV83507	Amphotrophic MLV r
33	308.8	10.4	2484	AAV05253	Amphotrophic MLV r
34	308.8	10.4	8440	AAV13265	Sequence of clone
35	307.6	10.4	1917	AAV89075	FelV F6A provirus
36	307.6	10.4	1998	AAV83509	Amphotrophic glyco
37	307.2	10.4	1602	AAV89139	Amphotrophic MLV r
38	307.2	10.4	1959	AAV89074	Canine retrovirus
39	307.2	10.4	6028	AAV90689	Amphotrophic MLV en
40	306.8	10.4	1944	AAV83508	Plasmid pBdelPASA
41	303.8	10.3	1965	AAV09276	Amphotrophic MLV r
42	302.2	10.2	1965	AAV83506	V/H recombinant pr
43	302.2	10.2	1950	AAV83505	Amphotrophic MLV r
44	301.8	10.2	1926	AAV89077	Amphotrophic MLV r
45	301.8	10.2	1929	AAV83503	Marine leukemia vi
					Amphotrophic MLV r

## ALIGNMENTS

AAV09699	1	AAV09699 standard; DNA: 8196 BP.
AAV09699		
19-MAY-1998		(first entry)
XX		Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX		Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;
KW		vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX		
OS		Porcine retrovirus.
XX		
FH	Key	Location/Qualifiers
FT	CDS	576..2126
FT		/*tag= a
FT		/product= GAG protein
FT	CDS	/note= "viral core polypeptide"
FT		2143..5733
FT		/*tag= b
FT		/product= POL protein
FT		/note= "polymerase peptide sequence as given in
FT		specification"
FT		5606..7576
FT		/*tag= c
FT		/product= ENV protein
FT		/note= "envelope protein"
PN		W09740167-A1.
XX		
PD		30-OCT-1997.
XX		

PF 18-APR-1997: 97WO-GB01087.  
 XX  
 PR 10-FEB-1997: 97GB-0002668.  
 PR 19-APR-1996: 96GB-0008164.  
 XX  
 PA (IMUT-) IMUTRAN LTD.  
 PA (OONE-) Q-ONE BIOTECH LTD.  
 PI Galbraith DN, Haworth C, Lees GM, Smith KT;  
 XX WPI, 1997-535851/49.  
 DR  
 XX Polynucleotide encoding porcine retrovirus expression product -  
 PT useful to develop products for use in vaccines, diagnosis and  
 PS xeno-transplantation  
 XX  
 PS Claim 4; Fig 2; 69pp; English.  
 CC This DNA sequence encodes the porcine retrovirus (PoEV) virion core  
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These  
 CC proteins can be used to develop viral vaccines, antisense nucleic acids,  
 CC ribozymes and other antiviral agents. They can also be used in  
 CC xeno-transplantation technology and as diagnostic tools.  
 XX  
 SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

Query Match 98.0%; Score 2896.8; DB 18; Length 8196;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 2937; Conservative 3; Mismatches 10; Indels 6; Gaps 3;

QY 1 tgccttttaaggttagaacaacccctggaacagtttggtgctgacccctatgattgctct 60  
 DB 5246 tgccttttaaggttagaacaacccctggaacagtttggtgctgacccctatgattgctct 5305  
 QY 61 acggggagaccccccgttgtagaattgctctctacatagtgctgattgctctt 120  
 DB 5306 acggggagaccccccgttgtagaattgctctctacatagtgctgattgctctt 5365  
 QY 121 cccagccttgctctcctagctcaaggcgtcagtggtgtagaagcaagcgctgagc 180  
 DB 5366 cccagccttgctctcctagctcaaggcgtcagtggtgtagaagcaagcgctgagc 5425  
 QY 181 agcccgaggagcctactcagagagagagacttgcaagttccacatcgcttccaaagt 240  
 DB 5426 agcccgaggagcctactcagagagagagacttgcaagttccacatcgcttccaaagt 5482  
 QY 241 gagaatcagctatgtagacgcacccgtgcaagaaaccccgagactcggtggaagagac 300  
 DB 5483 gagaatcagctatgtagacgcacccgtgcaagaaaccccgagactcggtggaagagac 5542  
 QY 301 ctatctcgtactttgacacacacaaaggctgtagaagtcgaaggaatctccactgga 360  
 DB 5543 ctatctcgtactttgacacacacaaaggctgtagaagtcgaaggaatctccactgga 5602  
 QY 361 tcatgtagatccacagtaagctgagcgccactcccgactgaggtgtagagcgcaagaaga 420  
 DB 5603 tcatgtagatccacagtaagctgagcgccactcccgactgaggtgtagagcgcaagaaga 5662  
 QY 421 ctggaatccoccttaagctcgccctcactcgctggtcttctactacttaaaactcc 480  
 DB 5663 ctggaatccoccttaagctcgccctcactcgctggtcttctactacttaaaactcc 5722  
 QY 481 caggcagtagtaaacgcttatagacagctgaaaccccatagaccttatccctacc 540  
 DB 5723 caggcagtagtaaacgcttatagacagctgaaaccccatagaccttatccctacc 5782  
 QY 541 tgggtgattatgacctgtagcgggtgtgtactgtaaatagcactgaggtgtgtctcc 600  
 DB 5783 tgggtgattatgacctgtagcgggtgtgtactgtaaatagcactgaggtgtgtctcc 5842  
 QY 601 agagacccctggtgctgactgacttctgctcctcgatgtaaaccccggtgtaaa 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 5843 agagacccctggtgctgactgacttctgctcctcgatgtaaaccccggtgtaaa 5902  
 QY 661 agcacacccctccaaacctgctgtagttaggtgtctatctgctccagcagagagaa 720  
 DB 5903 agcacacccctccaaacctgctgtagttaggtgtctatctgctccagcagagagaa 5962  
 QY 721 gagaatactgtggtggtctggtggaatcctctctgtagagatggaagctgcactcc 780  
 DB 5963 gagaatactgtggtggtctggtggaatcctctctgtagagatggaagctgcactcc 6022  
 QY 781 aacgatatgagactggaatactgctcctcctccagcagcggtgaataatctctctg 840  
 DB 6023 aacgatatgagactggaatactgctcctcctccagcagcggtgaataatctctctg 6082  
 QY 841 aattccggcccggaagtaacaaatgaagaactataaaagaataagagctgtcccca 900  
 DB 6083 aattccggcccggaagtaacaaatgaagaactataaaagaataagagctgtcccca 6142  
 QY 901 tgaacttagatattatcctaagaataagttcactgaaagaaacaggaataatccaa 960  
 DB 6143 tgaacttagatattatcctaagaataagttcactgaaagaaacaggaataatccaa 6200  
 QY 961 aagtgatataatggtatgagctggtggaatgttttaataataatggtggtggaag 1020  
 DB 6201 aagtgatataatggtatgagctggtggaatgtttttt-attatattggtggtggaag 6259  
 QY 1021 tccactttaaccatcgcttagatagatagacagggagacagaaacccctgtgcaatgga 1080  
 DB 6260 tccactttaaccatcgcttagatagatagacagggagacagaaacccctgtgcaatgga 6319  
 QY 1081 cccgataagatgtagctgctggaacaggggcccggccggtggaaccccgataactg 1140  
 DB 6320 cccgataagatgtagctgctggaacaggggcccggccggtggaaccccgataactg 6379  
 QY 1141 gtcgcccaatlaacccctgctgctgcaataacacacagccgcttagaacaagtaacca 1200  
 DB 6380 gtcgcccaatlaacccctgctgctgcaataacacacagccgcttagaacaagtaacca 6439  
 QY 1201 ggaattgattcctacacacacagcctagaacatcccaagtttctggttaagaaggaag 1260  
 DB 6440 ggaattgattcctacacacacagcctagaacatcccaagtttctggttaagaaggaag 6499  
 QY 1261 agactcctcagctcatccagagagcttccaaagcacaactccacagccagctatgcc 1320  
 DB 6500 agactcctcagctcatccagagagcttccaaagcacaactccacagccagctatgcc 6559  
 QY 1321 acttctctgtgtgctgtctatccctcagggcctcctatlatgagggatggtctaaa 1380  
 DB 6560 acttctctgtgtgctgtctatccctcagggcctcctatlatgagggatggtctaaa 6619  
 QY 1381 gaaagaataatcaatgtagcacaagaagatagaaatcaatgtagatggtgtccgaat 1440  
 DB 6620 gaaagaataatcaatgtagcacaagaagatagaaatcaatgtagatggtgtccgaat 6679  
 QY 1441 aagcttaacctcactggaatttcgggaagggagatgcatagagaaagctccccaacc 1500  
 DB 6680 aagcttaacctcactggaatttcgggaagggagatgcatagagaaagctccccaacc 6739  
 QY 1501 caccacaacacttgctatagactggtgtgtatgtagcagggcctcagaataatgatttta 1560  
 DB 6740 caccacaacacttgctatagactggtgtgtatgtagcagggcctcagaataatgatttta 6799  
 QY 1561 gtaactggttatcaaacaggtgtgtggaatgcaatctggtgttaacccctgtgtccacc 1620  
 DB 6800 gtaactggttatcaaacaggtgtgtggaatgcaatctggtgttttccacc 6859  
 QY 1621 tcaagcttaacaaatccaaagatttctgtgtcatggtgtccaaatggtcccgaggtgac 1680  
 DB 6860 tcaagcttaacaaatccaaagatttctgtgtcatggtgtccaaatggtcccgaggtgac 6919  
 QY 1681 taccatctgaaagatggtctctgtatgtaataatgactatcggtataaccgacaaagaa 1740  
 DB 6920 taccatctgaaagatggtctctgtatgtaataatgactatcggtataaccgacaaagaa 6979

QY 1741 gaaccgatacccttaacctagctgtatagctcggatlaaggagccgcttgagcgtaga 1800  
 |||||  
 Db 6980 gaaccgatacccttaacctagctgtatagctcggatlaaggagccgcttgagcgtaga 7039  
 QY 1801 acaaggagcagctgcccgtatcacagagccacagcagctagaagaagacttgtagcta 1860  
 |||||  
 Db 7040 acaaggagcagctgcccgtatcacagagccacagcagctagaagaagacttgtagcta 7099  
 QY 1861 catcgcccatcagacagaagatctccgagccttaaggagctgttaagcaacctagaagag 1920  
 |||||  
 Db 7100 catcgcccatcagacagaagatctccgagccttaaggagctgttaagcaacctagaagag 7159  
 QY 1921 tccctgacctcttgcttgagtggttctacagaaccgaggggattagatagctctgcttt 1980  
 |||||  
 Db 7160 tccctgacctcttgcttgagtggttctacagaaccgaggggattagatagctctgcttt 7219  
 QY 1981 ctaagagaagtggtgtatgtgcagccttaagaagaatgtgtctctatgtatagatcac 2040  
 |||||  
 Db 7220 ctaagagaagtggtgtatgtgcagccttaagaagaatgtgtctctatgtatagatcac 7279  
 QY 2041 tcaagagccatcagacagaactcattgaacaagcttagaagaagctagaagcgctgaag 2100  
 |||||  
 Db 7280 tcaagagccatcagacagaactcattgaacaagcttagaagaagctagaagcgctgaag 7339  
 QY 2101 gaaagagagctgaaccaggggtgttgtaagagtggttcacagctcctctgtagatgacc 2160  
 |||||  
 Db 7340 gaaagagagctgaaccaggggtgttgtaagagtggttcacagctcctctgtagatgacc 7399  
 QY 2161 accctgtctctgctcctgaaggggcccctagtagctcgtcctgtacttaacagttag 2220  
 |||||  
 Db 7400 accctgtctctgctcctgaaggggcccctagtagctcgtcctgtacttaacagttag 7459  
 QY 2221 ccttgcttaataagtgcttgcttgcttgcttgtagaagaagtgagtgagccagatc 2280  
 |||||  
 Db 7460 ccttgcttaataagtgcttgcttgcttgcttgtagaagaagtgagtgagccagatc 7519  
 QY 2281 atgttacttaggaacaagtaaccaagggcctctgagccaagaggaactgaacctagcct 2340  
 |||||  
 Db 7520 atgttacttaggaacaagtaaccaagggcctctgagccaagaggaactgaacctagcct 7579  
 QY 2341 tcccaagtcttaagaattagaacttlaacagaacaagaatgtgggaatgaagaatgaaga 2400  
 |||||  
 Db 7580 tcccaagtcttaagaattagaacttlaacagaacaagaatgtgggaatgaagaatgaaga 7639  
 QY 2401 tgaacctaaacctccacagaaccagaagatlaataaagctctaaatgccccgaatt 2460  
 |||||  
 Db 7640 tgaacctaaacctccacagaaccagaagatlaataaagctctaaatgccccgaatt 7699  
 QY 2461 acaagaccctgctgctgcagtaaatagtagaaggtcaactctctatgttccagggc 2520  
 |||||  
 Db 7700 mcaagaccctgctgctgcagtaaatagtagaaggtcaactctctatgttccagggc 7759  
 QY 2521 ctgctatccctgagctaagaatlaacagaagaatgagtggtgactaaatcgcttatctgatt 2580  
 |||||  
 Db 7760 ctgctatccctgagctaagaatlaacagaagaatgagtggtgactaaatcgcttatctgatt 7819  
 QY 2581 ctgtlaaactgacttgccacatagaagaattgattacacatgacagcccttagtgacct 2640  
 |||||  
 Db 7820 ctgtlaaactgacttgccacatagaagaattgattacacatgacagcccttagtgacct 7879  
 QY 2641 tctcaactgcaactctgctacatcgtccagagagccccaagatgagcagctcccgagcta 2700  
 |||||  
 Db 7880 tctcaactgcaactctgctacatcgtccagagagccccaagatgagcagctcccgagcta 7939  
 QY 2701 tttaaaatgagtgctccagagcgcggtcctgatattttaaatattgtgtccatg 2760  
 |||||  
 Db 7940 tttaaaatgagtgctccagagcgcggtcctcctgatattttaaatattgtgtccatg 7999  
 QY 2761 gaagcggggtcctgatattttaaaatgattgtttgtgacgacaggttctgttgtaa 2820  
 |||||  
 Db 8000 gaagcggggtcctcctgatattttaaaatgattgtttgtgacgacaggttctgttgtaa 8059

QY 2821 ccccatlaaagctgtcccgatcccgactcgggcccagctcctacccctgctgtg 2880  
 |||||  
 Db 8060 ccccatlaaagctgtcccgatcccgactcgggcccagctcctacccctgctgtg 8119  
 QY 2881 taagactgtgggcccagcgcttggaataaaatcctctgctgtttgtcatcaaaaa 2940  
 |||||  
 Db 8120 taagactgtgggcccagcgcttggaataaaatcctctgctgtttgtcatcaaaaa 8179  
 QY 2941 aaaaaaaaaaaaaa 2956  
 |||||  
 Db 8180 aaaaaaaaaaaaaa 8195  
 RESULT 2  
 AAV09700  
 ID AAV09700 standard; DNA; 8209 BP.  
 AC AAV09700;  
 XX  
 DT 20-MAY-1998 (first entry)  
 XX  
 DE Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.  
 XX  
 KW Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;  
 KW virion core polypeptide; polymerase protein; envelope protein;  
 KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
 XX  
 OS Porcine retrovirus.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT 1..61  
 FT misc\_feature  
 FT /tag= a  
 FT /note= "R LTR domain"  
 FT 1..588  
 FT LTR  
 FT /tag= b  
 FT 62..143  
 FT misc\_feature  
 FT /tag= c  
 FT /note= "U5 LTR domain"  
 FT 588..2162  
 FT CDS  
 FT /tag= d  
 FT /product= GAG protein  
 FT /note= "viral core polypeptide"  
 FT 2163..5747  
 FT CDS  
 FT /tag= e  
 FT /product= POL protein  
 FT /note= "polymerase peptide"  
 FT 5748..7590  
 FT CDS  
 FT /tag= f  
 FT /product= ENV protein  
 FT /note= "envelope protein"  
 FT 7591..8209  
 FT LTR  
 FT /tag= g  
 FT 7638..8106  
 FT misc\_feature  
 FT /tag= h  
 FT /note= "U3 LTR domain"  
 FT 7643..7686  
 FT protein\_bind  
 FT /tag= i  
 FT /note= "Binding site for transcription factors  
 FT OCT-1, c-Myb, Myf-1, E47 and ETS-1"  
 FT 7719..7739  
 FT protein\_bind  
 FT /tag= j  
 FT /note= "Binding site for transcription factors  
 FT AP-4 and NF1"  
 FT 7744..7754  
 FT protein\_bind  
 FT /tag= k  
 FT /note= "Binding site for transcription factor AP-1/TR"  
 FT 7773..7785  
 FT protein\_bind  
 FT /tag= l  
 FT /note= "Binding site for transcription factor ETS-1/GATA"  
 FT 7793..7834  
 FT protein\_bind  
 FT /tag= m  
 FT /note= "Binding site for transcription factors  
 FT GATA, ETS-1, c-Myb, AP-1, GATA, E47"  
 FT

FT protein\_bind 7841..7851  
 FT /tag= n  
 FT /note= "Binding site for transcription factor AP-1"  
 FT 7873..7883  
 FT /tag= o  
 FT /note= "Binding site for transcription factor AP-1"  
 FT 7885..7914  
 FT /tag= p  
 FT /note= "Binding site for transcription factors  
 FT AP-1/GATA, c-Myb and AP-1"  
 FT 7928..7953  
 FT /tag= q  
 FT /note= "Binding site for transcription factors E47  
 FT and ETS-1"  
 FT 7964..7968  
 FT /tag= r  
 FT 7985..7997  
 FT /tag= s  
 FT /note= "Binding site for transcription factor GATA"  
 FT 8003..8007  
 FT /tag= t  
 FT 8024..8036  
 FT /tag= u  
 FT /note= "Binding site for transcription factor GATA"  
 FT 8042..8046  
 FT /tag= v  
 FT 8049..8060  
 FT /tag= w  
 FT /note= "Binding site for transcription factor  
 FT TATA\_signal 8075..8090  
 FT /tag= x  
 FT 1..61  
 FT /tag= y  
 FT /note= "R LRR domain"  
 FT 8127..8135  
 FT /tag= z  
 FT /note= "PolyA downstream element"  
 FT 8161..8166  
 FT /tag= aa  
 PN WO9740167-A1.  
 PD 30-OCT-1997.  
 XX 18-APR-1997: 97WO-GB01087.  
 XX 10-FEB-1997: 97GB-0002668.  
 PR 19-APR-1996: 96GB-0008164.  
 PA (IMUT-) IMUTRAN LTD.  
 PA (OONE-) O-ONE BIOTEC LTD.  
 PI Galbraith DN, Haworth C, Lees GM, Smith KT;  
 DR WPI: 1997-535851/49.  
 DR P-PsDB: AAM39271; AAM39272; AAM39273.  
 XX Polynucleotide encoding porcine retrovirus expression product -  
 PT useful to develop products for use in vaccines, diagnosis and  
 PT xeno-transplantation  
 XX Claim 4; Fig 3; 69pp: English.  
 CC This DNA sequence encodes the porcine retrovirus (POEV) virion core  
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and  
 CC also includes the long terminal Repeat (LTR). These proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.  
 XX Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

Query Match 98.0%; Score 2896.4; DB 18; Length 8209;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 2939; Conservative 0; Mismatches 11; Indels 6; Gaps 3;  
 QY 1 tgcctttagaggttagaacacccttgacaggttggctgtaacccctatgaattgctc 60  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5260 tgcctttagaggttagaacacccttgacaggttggctgtaacccctatgaattgctc 5319  
 QY 61 aagggggagcccccccgctgtgtagaattgctctcctgacatagtcgtatgctgctt 120  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5320 aagggggagcccccccgctgtgtagaattgctctcctgacatagtcgtatgctgctt 5379  
 QY 121 cccagccttgttctcctagctcaagcgctcgaagtgtgtgaggaacagcgtggaagc 180  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5380 cccagccttgttctcctagctcaagcgctcgaagtgtgtgaggaacagcgtggaagc 5439  
 QY 181 agctccggagagccttaccagagagagagagagagagagagagagagagagagag 240  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5440 agctccggagagccttaccagagagagagagagagagagagagagagagagagag 5496  
 QY 241 gagatcagctatgttagacgcacacgctgacggaacacacacacacacacacacac 300  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5497 gagatcagctatgttagacgcacacgctgacggaacacacacacacacacacacac 5556  
 QY 301 ctatctcgtactttgac 360  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5557 ctatctcgtactttgac 5616  
 QY 361 tccatgcattccacaggttagacgcacacacacacacacacacacacacacacacacac 420  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5617 tccatgcattccacaggttagacgcacacacacacacacacacacacacacacacacac 5676  
 QY 421 ctgagatcccttaagctcgcctccacacacacacacacacacacacacacacacacac 480  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5677 ctgagatcccttaagctcgcctccacacacacacacacacacacacacacacacacac 5736  
 QY 481 cagggcagtagtaaacgccttatagacagctcgaaccccaatagaaccttatccctacc 540  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5737 cagggcagtagtaaacgccttatagacagctcgaaccccaatagaaccttatccctacc 5796  
 QY 541 tggcgtattatggacccctgacaggtgtacatgtgaatgaacacacacacacacacac 600  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5797 tggcgtattatggacccctgacaggtgtacatgtgaatgaacacacacacacacacac 5856  
 QY 601 agaggcacctgtgtgctcgaactgcatcttgcctccacacacacacacacacacacac 660  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5857 agaggcacctgtgtgctcgaactgcatcttgcctccacacacacacacacacacacac 5916  
 QY 661 agcacactcccaactatgctcgtatgtgttctatctgtcgtccagacagagagaa 720  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5917 agcacactcccaactatgctcgtatgtgttctatctgtcgtccagacagagagaa 5976  
 QY 721 gagaatactgtgggggttcctggggaacccctcgttaagaatgaagctgtgcaacctc 780  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 5977 gagaatactgtgggggttcctggggaacccctcgttaagaatgaagctgtgcaacctc 6036  
 QY 781 aacgattgagactggaatgtgcgaatctctctccagacagcggtlaaaatctctctgtc 840  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 6037 aacgattgagactggaatgtgcgaatctctctccagacagcggtlaaaatctctctgtc 6096  
 QY 841 aattccggcccgagggaatgataaaatgtatgaatactatgaagatgaagctgtcccca 900  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 6097 aattccggcccgagggaatgataaaatgtatgaatactatgaagatgaagctgtcccca 6156  
 QY 901 tcaagcttagattatcctaagaataaagttcctcgtgaagaagaagaagaagaatcaca 960  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 6157 tcaagcttagattatcctaagaataaagttcctcgtgaagaagaagaagaagaatcaca 6214  
 QY 961 aagtgtataaatgtatagctgtgggaatgttttaataatatgtgcggggagcagg 1020  
 DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 DB 6215 aagtgtataaatgtatagctgtgggaatgtttt-attatagtgggggagcagg 6273

QY 1021 tccatttaaccattcgccttagatagagacgaggacagaaacccctgtgccaatggga 1080  
Db 6274 tccatttaaccattcgccttagatagagacgaggacagaaacccctgtgccaatggga 6333  
QY 1081 cccgataaagctacgtgctggaacagggccccgcgccttgagaccacgcataactgtgcg 1140  
Db 6334 cccgataaagctacgtggaacagggccccgcgccttgagaccacgcataactgtgcg 6393  
QY 1141 gtgcaccaatlaacctgcgtgcgcgtgcatacaacagccgccttagacaagctaacct 1200  
Db 6394 gtgcaccaatlaacctgcgtgcgcgtgcatacaacagccgccttagacaagctaacct 6453  
QY 1201 ggattgtatccctaccaaacgcgtagaaactcccaggtgttccgtgttaagacagagag 1260  
Db 6454 ggattgtatccctaccaaacgcgtagaaactcccaggtgttccgtgttaagacagagag 6513  
QY 1261 agactctcagctcattccacagggagcttccacgcataactccacgcacctgtatgcc 1320  
Db 6514 agactctcagctcattccacagggagcttccacgcataactccacgcacctgtatgcc 6573  
QY 1321 actctctctgtgtgcttcttctatccctcagggccctctattatgagggatgtgctaa 1380  
Db 6574 actctctctgtgtgcttcttctatccctcagggccctctattatgagggatgtgctaa 6633  
QY 1381 gaaagaaatccaatgtgaccaaagagcatalaataatgtacatgtgggtlcccgaaat 1440  
Db 6634 gaaagaaatccaatgtgaccaaagagcatalaataatgtacatgtgggtlcccgaaat 6693  
QY 1441 aagcttaccctcactgtaagttccgggaagggacatgcatagaaagctcccccatacc 1500  
Db 6694 aagcttaccctcactgtaagttccgggaagggacatgcatagaaagctcccccatacc 6753  
QY 1501 caccacaacctgtgctatgtaactgtgttctatgagcagccctcagaataatcgatitaa 1560  
Db 6754 caccacaacctgtgctatgtaactgtgttctatgagcagccctcagaataatcgatitaa 6813  
QY 1561 gtaccgtgttataacagagtggtgacatgcataactggtgttaacccctgtgttccacc 1620  
Db 6814 gtaccgtgttataacagagtggtgacatgcataactggtgtgttaacccctgtgttccacc 6873  
QY 1621 tcaagtctcaaccaatccaagaattgtgtgtcatgtgtccaatactgcccgaggtac 1680  
Db 6874 tcaagtctcaaccaatccaagaattgtgtgtcatgtgtccaatactgcccgaggtac 6933  
QY 1681 taaccatctgaggaagtggtctctgtgaatgaatgactatcgatataaccgacaataaaga 1740  
Db 6934 taaccatctgaggaagtggtctctgtgaatgaatgactatcgatataaccgacaataaaga 6993  
QY 1741 gaaccgcgtatcccttaaccctgaagctgtgaatgtcgtgattaggagcggttggtcgtaga 1800  
Db 6994 gaaccgcgtatcccttaaccctgaagctgtgaatgtcgtgattaggagcggttggtcgtaga 7053  
QY 1801 aacagggaacagctgcctctgatacagagccacagcagctcagagaaagagactgtgtgcgta 1860  
Db 7054 aacagggaacagctgcctctgatacagagccacagcagctcagagaaagagactgtgtgcgta 7113  
QY 1861 catgctggccaatgacagaaatctccgagccttaaaagagtgctgttagcaactcagagagag 1920  
Db 7114 catgctggccaatgacagaaatctccgagccttaaaagagtgctgttagcaactcagagagag 7173  
QY 1921 tccctgaattcttctgtcgtgaagtggtctacagaaacggaggggattagatctgctgtc 1980  
Db 7174 tccctgaattcttctgtcgtgaagtggtctacagaaacggaggggattagatctgctgtc 7233  
QY 1981 ctaaggaagtggtgtatgtgacgaccttaaaagaaatgtgtgctctcatgtatgatacc 2040  
Db 7234 ctaaggaagtggtgtgtatgtgacgaccttaaaagaaatgtgtgctctcatgtatgatacc 7293  
QY 2041 tcaaggagccatcagagactccatgaaacagcttagaanaaagttagagagggcgtcgaag 2100  
Db 7294 tcaaggagccatcagagactccatgaaacagcttagaanaaagttagagagggcgtcgaag 7353  
QY 2101 gaaagagaggtcagcaccaggggtgtgttagagatgtgtcaacaggtctcctcgatgacc 2160

Db 7354 gaaagagaggtcagcaccaggggtgtgttagagatgtgtcaacaggtctcctcgatgacc 7413  
QY 2161 accctgtcttctgtctgtacagggccccctagtagtccctgtcctctgattactaacgttggg 2220  
Db 7414 accctgtcttctgtctgtacagggccccctagtagtccctgtcctctgattactaacgttggg 7473  
QY 2221 ccttgcttaattatagtggttctgtccttctgttagagaaacaggtgagatccaagatc 2280  
Db 7474 ccttgcttaattatagtggttctgtccttctgttagagaaacaggtgagatccaagatc 7533  
QY 2281 atgttacttaggaaacagtaaccaagagctctcgaagcacaaggagaaactactcctagcct 2340  
Db 7534 atgttacttaggaaacagtaaccaagagctctcgaagcacaaggagaaactactcctagcct 7593  
QY 2341 tcccagttctaagattagaactaatacaagaacaaagtggtggaatgaaagatgaa 2400  
Db 7594 tcccagttctaagattagaactaatacaagaacaaagtggtggaatgaaagatgaa 7653  
QY 2401 tgaaccttaacctccccaagcccggaagttaataaagctcttaaatgccccgaatt 2460  
Db 7654 tgaaccttaacctccccaagcccggaagttaataaagctcttaaatgccccgaatt 7713  
QY 2461 acagacctgtcgtgctgcagtaaatagtgaaaggtcacaacttctatgttccagggc 2520  
Db 7714 acagacctgtcgtgctgcagtaaatagtgaaaggtcacaacttctatgttccagggc 7773  
QY 2521 ctgtctatccctgacctaaagtaagaatacagaaatgagttgacttaactcgtatcgtat 2580  
Db 7774 ctgtctatccctgacctaaagtaagaatacagaaatgagttgacttaactcgtatcgtat 7833  
QY 2581 ctgttaaaactgtactgccaacataagaagaattgattacacatgtagacgccttagtacta 2640  
Db 7834 ctgttaaaactgtactgccaacataagaagaattgattacacatgtagacgccttagtacta 7893  
QY 2641 tctcaactgcaatctgtcaactctgtccaggaagccacagacagatgtggaactcccgagcta 2700  
Db 7894 tctcaactgcaatctgtcaactctgtccaggaagccacagacagatgtggaactcccgagcta 7953  
QY 2701 ttcttaaaatgattgtgtccacggagcgcggtcctcgatatcttaataatgattgtccatg 2760  
Db 7954 ttcttaaaatgattgtgtgtccacggagcgcggtcctcgatatcttaataatgattgtccatg 8013  
QY 2761 gagcgcggtcctcgatatcttaataatgattgtgtgtgacgacagacgctgtgtgtaa 2820  
Db 8014 gagcgcggtcctcgatatcttaataatgattgtgtgtgacgacagacgctgtgtgtaa 8073  
QY 2821 ccccatlaaagctgtcccgattccgactccgggcggcgagctcctacccctgctgtg 2880  
Db 8074 ccccatlaaagctgtcccgattccgactccgggcggcgagctcctacccctgctgtg 8133  
QY 2881 taagactgtggcccccaagcgcgcttggaataaaatcctctgtcgtttagcataaaaaa 2940  
Db 8134 taagactgtggcccccaagcgcgcttggaataaaatcctctgtcgtttagcataaaaaa 8193  
QY 2941 aaaaaaaaaaaaaa 2956  
Db 8194 aaaaaaaaaaaaaa 8209

RESULT 3  
AAV82749  
ID AAV82749 standard; DNA; 3482 BP.  
XX  
XX AAV82749;  
AC  
DT 25-FEB-1999 (first entry)  
XX  
XX  
DE Pig endogenous retrovirus (PERV)-B envelope (env) gene region.  
KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
KW probe; detection; retrovirus; human tissue; xenotransplant;  
KW primary porcine tissue; human cell line; porcine cell line; ss.

XX OS Pig endogenous retrovirus.  
 XX FH Key Location/Qualifiers  
 XX CDS 911..2884  
 FT /+tag= a  
 FT /product= envelope\_protein  
 XX PN WO9853104-A2.  
 XX PD 26-NOV-1998.  
 XX PF 18-MAY-1998; 98MO-GB01428.  
 XX PR 16-MAY-1997; 97GB-0010154.  
 XX PA (MEDI-) MEDICAL RES COUNCIL.  
 XX PI Stoye JP, Weiss RA;  
 XX WPI: 1999-045324/04.  
 DR P-PSDB: AAW85452.  
 XX  
 PT Newly isolated nucleic acid probe capable of hybridising to either  
 PT the PERV-A or PERV-B env gene, useful in the detection of  
 PT retroviruses, and their subtypes, in a sample of porcine/human  
 PT tissue  
 PS Claim 3: Page 21-23; 36pp; English.  
 XX  
 CC The present sequence encodes a Pig endogenous retrovirus (PERV)-B  
 CC envelope protein. PERV exists in two different subtypes, PERV-A and  
 CC PERV-B. The differences are reflected in sequence divergence in the  
 CC envelope genes. Probes and primers can be derived from the envelope  
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
 CC a method to detect retroviruses in a sample of porcine/human tissue.  
 CC particularly primary porcine tissue and human cell lines that have been  
 CC cultivated in the presence of a porcine cell line, or human tissue from  
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
 CC one of the PERV env genes can also be determined.  
 CC  
 SO Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

Query Match 97.7%; Score 2889.4; DB 20; Length 3482;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 2913; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 1 tgccttttaaggttagaacaacccctggaacagtttggtgaccccttaaatgtctt 60  
 DB 548 tgccttttaaggttagaacaacccctggaacagtttggtgaccccttaaatgtctt 607  
 QY 61 acgggggaaccccccgttgtagaattgtctctgtatagtgctgtgtctctt 120  
 DB 608 acgggggaaccccccgttgtagaattgtctctgtatagtgctgtgtctctt 667  
 QY 121 ccaagccttgtctcttaggtctaaagcgtcagtggtgtagaagcaagcgtgtgaac 180  
 DB 668 ccaagccttgtctcttaggtctaaagcgtcagtggtgtagaagcaagcgtgtgaac 727  
 QY 181 agctccggagagcctactcagagagagagacttgcaagtccacatcctcccaagt 240  
 DB 728 agctccggagagcctactcagagagagagacttgcaagtccacatcctcccaagt 787  
 QY 241 gaagattcagctatgttagaagccacgcgtgcaagaaactcgagactcggtgaagggac 300  
 DB 788 gaagattcagctatgttagaagccacgcgtgcaagaaactcgagactcggtgaagggac 847  
 QY 301 ctatctcgtacttttgaccacacaaagcgtgtgaagtgaagaaatcccaacttga 360  
 DB 848 ctatctcgtacttttgaccacacaaagcgtgtgaagtgaagaaatcccaacttga 907  
 QY 361 tccatgcatcccaagcttaagctgtgagccacactccgcagccgggtgtagagagccgaaaga 420

DB 908 tccatgcatcccaagcttaagctgtgagccacactccgcagctcgggtgtagagccgaaaga 967  
 QY 421 ctgagaatcccttaagcttcgctccatcgccctgtgtcttaacttaataactatcc 480  
 DB 968 ctgagaatcccttaagcttcgctccatcgccctgtgtcttaacttaataactatcc 1027  
 QY 481 cagccagtagtaaacgcttataagacagctcgaaccccaatagacttaccctacc 540  
 DB 1028 cagccagtagtaaacgcttataagacagctcgaaccccaatagacttaccctacc 1087  
 QY 541 tgcctgattatgtgacccctgatacgggtgtcaactgtaaatagacactcgaggtgtgtcct 600  
 DB 1088 tgcctgattatgtgacccctgatacgggtgtcaactgtaaatagacactcgaggtgtgtcct 1147  
 QY 601 agagcactctgtgctggaacttgatcttcgctccgatttgatgaaccccggtttaa 660  
 DB 1148 agagcactctgtgctggaacttgatcttcgctccgatttgatgaaccccggtttaa 1207  
 QY 661 agcacactcccaactagtcctcgtatgtatgtgtctatgtctgcccagacagagaaa 720  
 DB 1208 agcacactcccaactagtcctcgtatgtatgtgtctatgtctgcccagacagagaaa 1267  
 QY 721 gagaaatctgtggtgtctggtggaatcctctgtgtagaagatgtagctgtgtaacctcc 780  
 DB 1268 gagaaatctgtggtgtctggtggaatcctctgtgtagaagatgtagctgtgtaacctcc 1327  
 QY 781 aagctatgagactggaatctgagcctcctccagacccgggtaaaatctcctttgtc 840  
 DB 1328 aagctatgagactggaatctgagcctcctccagacccgggtaaaatctcctttgtc 1387  
 QY 841 aattccgcccggcgaaglacaaaatgtatgaactataataagaatagactgtcccca 900  
 DB 1388 aattccgcccggcgaaglacaaaatgtatgaactataataagaatagactgtcccca 1447  
 QY 901 tcaagcttagattacttaagaatagatttcctgtaaaaaaggaacccggaatattcaa 960  
 DB 1448 tcaagcttagattacttaagaatagatttcctgtaaaaaaggaacccggaatattcaa 1507  
 QY 961 aagtgataaatgtgataagctggtggaatagttttaaataataatggtggtggtgagcagg 1020  
 DB 1508 aagtgataaatgtgataagctggtggaatagttttaaataataatggtggtggtgagcagg 1567  
 QY 1021 tccacttaaccatcgcttaagatagagaagcgggagacacccctgtgtgcaatggga 1080  
 DB 1568 tccacttaaccatcgcttaagatagagaagcgggagacacccctgtgtgcaatggga 1627  
 QY 1081 cccgataaagtaactgtgctgaacaggggcccgcggccctggagccacgcataacttgcg 1140  
 DB 1628 cccgataaagtaactgtgctgaacaggggcccgcggccctggagccacgcataacttgcg 1687  
 QY 1141 gtcgcccaattaaactctgctggtgacacaaacacagccgcttaggaacagtaccact 1200  
 DB 1688 gtcgcccaattaaactctgctggtgacacaaacacagccgcttaggaacagtaccact 1747  
 QY 1201 ggtattgattcttaacaaagccttagaaaccccccaggtgtctctgtttaaagacagagag 1260  
 DB 1748 ggtattgattcttaacaaagccttagaaaccccccaggtgtctctgtttaaagacagagag 1807  
 QY 1261 agactctcagctcatccagagagacttcccaagccatcaactccacagacccgtatgac 1320  
 DB 1808 agactctcagctcatccagagagacttcccaagccatcaactccacagacccgtatgac 1867  
 QY 1321 actctctctgtgtgtctatccctcagggcctccttaataatgagggatgagctaa 1380  
 DB 1868 actctctctgtgtgtctatccctcagggcctccttaataatgagggatgagctaa 1927  
 QY 1381 gaaagaataatcaatgtgacaaagagatagaaatcaatgtatgaggggttccgaaat 1440  
 DB 1928 gaaagaataatcaatgtgacaaagagatagaaatcaatgtatgaggggttccgaaat 1987  
 QY 1441 aagctacccactgaagcttccggagagggacatgcataggaagaagctcccccctcc 1500



Db 1968 aagcttaccctcactgaagtttccgggaaggagacalgcagtagaanaagctcccccacc 2047  
 Qy 1501 caccacacacttgcctagtagtgcgtttatagcagcgccctagaanaatcagtatatta 1560  
 Db 2048 caccacacacttgcctagtagtgcgtttatagcagcgccctagaanaatcagtatatta 2107  
 Qy 1561 gtactcgtttataacaggttggcgcagtcagtaactcgtgttaacccctgtgtttccacc 1620  
 Db 2108 gtactcgtttataacaggttggcgcagtcagtaactcgtgttaacccctgtgtttccacc 2167  
 Qy 1621 tcaagttctcaacccaatccaagaatttgcgttcagtcagtcagtcagtcagtcagtcagtc 1680  
 Db 2168 tcaagttctcaacccaatccaagaatttgcgttcagtcagtcagtcagtcagtcagtcagtc 2227  
 Qy 1681 tacatccctgaggaagtgctcctgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1740  
 Db 2228 tacatccctgaggaagtgctcctgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2287  
 Qy 1741 gaaccggtatcccttaccctagctgcgttaagtcgcgttaagtcgcgttaagtcgcgtta 1800  
 Db 2288 gaaccggtatcccttaccctagctgcgttaagtcgcgttaagtcgcgttaagtcgcgtta 2347  
 Qy 1801 acagggagagctgcgcctgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1860  
 Db 2348 acagggagagctgcgcctgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2407  
 Qy 1861 catgcgcgcacatgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1920  
 Db 2408 catgcgcgcacatgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2467  
 Qy 1921 tccctgcagcttctgtcgtgaagtcgttcacagaaaccgagggagtagatcgtcgttt 1980  
 Db 2468 tccctgcagcttctgtcgtgaagtcgttcacagaaaccgagggagtagatcgtcgttt 2527  
 Qy 1981 ctaagagaaggttggttatgtgcagcctttaaagaagaatgtgtcctccttgcagtcagtc 2040  
 Db 2528 ctaagagaaggttggttatgtgcagcctttaaagaagaatgtgtcctccttgcagtcagtc 2587  
 Qy 2041 tcaagagccatcagagatccatgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2100  
 Db 2588 tcaagagccatcagagatccatgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2647  
 Qy 2101 gaaagagaagctgcagcaggggtgtgttgaaagatgtgtcaacagtcctcctgcagtcagtc 2160  
 Db 2648 gaaagagaagctgcagcaggggtgtgttgaaagatgtgtcaacagtcctcctgcagtcagtc 2707  
 Qy 2161 accctgtcttcgtcgtcagcaggggcccctagtcagtcagtcagtcagtcagtcagtcagtc 2220  
 Db 2708 accctgtcttcgtcgtcagcaggggcccctagtcagtcagtcagtcagtcagtcagtcagtc 2767  
 Qy 2221 ccttgtcttaataagtagtctgtgtcctcctgttagagaagagtagtcagtcagtcagtc 2280  
 Db 2768 ccttgtcttaataagtagtctgtgtcctcctgttagagaagagtagtcagtcagtcagtc 2827  
 Qy 2281 atgtgtacttgagcaacagtagcaagcctcctgcagccaagagaaactgcagtcagtc 2340  
 Db 2828 atgtgtacttgagcaacagtagcaagcctcctgcagccaagagaaactgcagtcagtc 2887  
 Qy 2341 tcccaagttcaagattagaactatgaacagacaagaagtgagggaatgaaagatgaataa 2400  
 Db 2888 tcccaagttcaagattagaactatgaacagacaagaagtgagggaatgaaagatgaataa 2947  
 Qy 2401 tgcacaactaaacctccagagaccaggaagtataaataagccttaaatgcagccgaatt 2460  
 Db 2948 tgcacaactaaacctccagagaccaggaagtataaataagccttaaatgcagccgaatt 3007  
 Qy 2461 acagagacctgtgcgtgcagtaataagtagaaggtcacactcctattgttcagggc 2520  
 Db 3008 ccagagacctgtgcgtgcagtaataagtagaaggtcacactcctattgttcagggc 3067  
 Qy 2521 ctgtatccctggcctaaagtagaataacgaagaatgtgtgactaacgcttctgtgatt 2580  
 Db 3068 ctgtatccctggcctaaagtagaataacgaagaatgtgtgactaacgcttctgtgatt 3127

Qy 2581 ctgtaaactgactggcaccatagagaagattgattacacattgcagccctagtcagtc 2640  
 Db 3128 ctgtaaactgactggcaccatagagaagattgattacacattgcagccctagtcagtc 3187  
 Qy 2641 tctcaactgcaactcgtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 2700  
 Db 3188 tctcaactgcaactcgtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 3247  
 Qy 2701 ttcttaaaatgattgtgtccagagagcgagcgtcctgcagatatttaaaatgattgttcac 2760  
 Db 3248 ttcttaaaatgattgtgtccagagagcgagcgtcctgcagatatttaaaatgattgttcac 3307  
 Qy 2761 gagcgcgagtc- tgcagatttttaaaatgattgtgtgtgcagcagagcgttgcagtc 2819  
 Db 3308 gagcgcgagtc- tgcagatttttaaaatgattgtgtgtgcagcagagcgttgcagtc 3367  
 Qy 2820 acccctaataagctgtccagatccgactcgcagtcgcagtcgcagtcagtcagtcagtc 2879  
 Db 3368 acccctaataagctgtccagatccgactcgcagtcgcagtcgcagtcagtcagtcagtc 3427  
 Qy 2880 gtacgactgtggcccccagcgagcgttggaataaaatcctctgtgtgttcac 2934  
 Db 3428 gtacgactgtggcccccagcgagcgttggaataaaatcctctgtgtgttcac 3482

RESULT 4  
 AAV09703  
 ID AAV09703 standard; DNA; 1974 BP.  
 XX  
 AC AAV09703;  
 XX  
 AC AAV09703 (first entry)  
 DT  
 DT 20-MAY-1998  
 XX  
 DE Porcine retrovirus Raji clone ENV DNA.  
 XX  
 KW Porcine retrovirus; POEV; ENV protein; envelope protein; vaccine;  
 KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
 XX  
 OS Porcine retrovirus.  
 XX  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 1..1974  
 FT /\*tag= a  
 FT /product= ENV protein  
 FT /note= "envelope protein"  
 FT  
 PN WO9740167-A1.  
 PD 30-OCT-1997.  
 XX  
 PD 18-APR-1997; 97WO-GB01087.  
 PF  
 PR 10-FEB-1997; 97GB-0002668.  
 PR 19-APR-1996; 96GB-0008164.  
 XX  
 PA (IMUT-) IMUTRAN LTD.  
 PA (QONE-) Q-ONE BIOTECH LTD.  
 XX  
 PI Galbraith DN, Haworth C, Lees GW, Smith KT;  
 DR WPI, 1997-535851/49.  
 DR P-PSDB; AAM39274.  
 XX  
 PT Polynucleotide encoding porcine retrovirus expression product -  
 PT useful to develop products for use in vaccines, diagnosis and  
 PT xeno-transplantation  
 PS  
 PS Claim 6; Fig 4; 6pp; English.  
 CC This sequence encodes the porcine retrovirus envelope (ENV) protein  
 CC isolated from the human cell line Raji. Such viral proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other

CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.

XX Sequence 1974 BP; 534 A; 489 C; 483 G; 468 T; 0 other;

Query Match 66.7%; Score 1970.8; DB 18; Length 1974;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1972; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 364 atgacccccaagttaagctgagccacccacccactgcgggtgagagcggaagaactg 423  
 |||||  
 Db 1 atgacccccaagttaagctgagccacccacccactgcgggtgagagcggaagaactg 60  
 QY 424 agaatcccttaagctgcgcctcactgcgtgttctctacttaacaataactcccg 483  
 |||||  
 Db 61 agaatcccttaagctgcgcctcactgcgtgttctctacttaacaataactcccg 120  
 QY 484 gccagtagtaaacgcttataagacagctcgaaaccccatagaccttatccctactcg 543  
 |||||  
 Db 121 gccagtagtaaacgcttataagacagctcgaaaccccatagaccttatccctactcg 180  
 QY 544 ctgattatgacccctgaacggtgtgtcaactgtlaaataagacacgcaggtgttgcctaga 603  
 |||||  
 Db 181 ctgattatgacccctgaacggtgtgtcaactgtlaaataagacacgcaggtgttgcctaga 240  
 QY 604 ggcacccgtggtgcctgaacgcttctgtcctcgcgtatgatacccggtgttaaaagc 663  
 |||||  
 Db 241 ggcacccgtggtgcctgaacgcttctgtcctcgcgtatgatacccggtgttaaaagc 300  
 QY 664 aacactcccaacccctagctcgtatgattatggttctatgtcgcgcagacagacagaaagag 723  
 |||||  
 Db 301 aacactcccaacccctagctcgtatgattatggttctatgtcgcgcagacagacagaaagag 360  
 QY 724 aaataactgtggtgttctgtgggaatccctctgtatagagatgagagctgtgcaccccaac 783  
 |||||  
 Db 361 aaataactgtggtgttctgtgggaatccctctgtatagagatgagagctgtgcaccccaac 420  
 QY 784 gatgagaactgtgaatgagccatctctctcagaacgggttaaaatctctctgttcaat 843  
 |||||  
 Db 421 gatgagaactgtgaatgagccatctctctcagaacgggttaaaatctctctgttcaat 480  
 QY 844 tccggtcccggtgaagtaacaaatgaaatataaagaataagagctgtccccaatca 903  
 |||||  
 Db 481 tccggtcccggtgaagtaacaaatgaaatataaagaataagagctgtccccaatca 540  
 QY 904 gactatagattactaaagataaacttcaactgaaaaagaaacgggaaatatcaaaag 963  
 |||||  
 Db 541 gactatagattactaaagataaacttcaactgaaaaagaaacgggaaatatcaaaag 600  
 QY 964 tggataaactgtatgagactggtggaatagtttttaataataatggcgggtgagcagggtcc 1023  
 |||||  
 Db 601 tggataaactgtatgagactggtggaatagtttttaataataatggcgggtgagcagggtcc 660  
 QY 1024 actttaaccatctgccttagatagatagacgggtgagacaaacccctgtgccaatgggaccc 1083  
 |||||  
 Db 661 actttaaccatctgccttagatagatagacgggtgagacaaacccctgtgccaatgggaccc 720  
 QY 1084 gataaactgactggtgacaaagggcccggtggtgagacaaacggcactaacttcgggtg 1143  
 |||||  
 Db 721 gataaactgactggtgacaaagggcccggtggtgagacaaacggcactaacttcgggtg 780  
 QY 1144 cccaatlaactgcgtgagcgcctgacataaacaacagcgcttagcaacagtaaccactgga 1203  
 |||||  
 Db 781 cccaatlaactgcgtgagcgcctgacataaacaacagcgcttagcaacagtaaccactgga 840  
 QY 1204 ttgattctactacaacagccttagaacttcccaaggtgttctgttgaagacagagcaaga 1263  
 |||||  
 Db 841 ttgattctactacaacagccttagaacttcccaaggtgttctgttgaagacagagcaaga 900  
 QY 1264 ctcttcaagctcacaacagaggttccaaagcactcaactccacggacccctgagccact 1323  
 |||||  
 Db 901 ctcttcaagctcacaacagaggttccaaagcactcaactccacggacccctgagccact 960

QY 1324 tctctgtgtgcttgcctatcctcagggccctcttataatgaaggtatgctaaagaa 1383  
 |||||  
 Db 961 tctctgtgtgcttgcctatcctcagggccctcttataatgaaggtatgctaaagaa 1020  
 QY 1384 agaaatlaactgtgaccaaagagatagaaatcaatgataatggtgtccgaataag 1443  
 |||||  
 Db 1021 agaaatlaactgtgaccaaagagatagaaatcaatgataatggtgtccgaataag 1080  
 QY 1444 cttaaccttaacgaagttccgggaagggagatcataggaagaagctcccatccac 1503  
 |||||  
 Db 1081 cttaaccttaacgaagttccgggaagggagatcataggaagaagctcccatccac 1140  
 QY 1504 caacaccttgcatactgactgtgttataagacggcctcagaanaatcagttatgta 1563  
 |||||  
 Db 1141 caacaccttgcatactgactgtgttataagacggcctcagaanaatcagttatgta 1200  
 QY 1564 cctgtgtataacaggtgtgtggcaatgcaatctggttaacccctgtgttccactca 1623  
 |||||  
 Db 1201 cctgtgtataacaggtgtgtggcaatgcaatctggttaacccctgtgttccactca 1260  
 QY 1624 gtcttcaaccatccaagattgtgtgtatgtgtccaaatcgttcccgaggtactac 1683  
 |||||  
 Db 1261 gtcttcaaccatccaagattgtgtgtatgtgtccaaatcgttcccgaggtactac 1320  
 QY 1684 catcctgaaaggtgtccttgatgataatgactatcggtataaccgaccaaagagaa 1743  
 |||||  
 Db 1321 catcctgaaaggtgtccttgatgataatgactatcggtataaccgaccaaagagaa 1380  
 QY 1744 cccgtatcccttaacccctagctgtatgttcgataatgaagagcggtgtgtgtagaaca 1803  
 |||||  
 Db 1381 cccgtatcccttaacccctagctgtatgttcgataatgaagagcggtgtgtgtagaaca 1440  
 QY 1804 gggacagctgtccctgatacacaagacacagcgcctagaagaaagactgtgtgtactac 1863  
 |||||  
 Db 1441 gggacagctgtccctgatacacaagacacagcgcctagaagaaagactgtgtgtactac 1500  
 QY 1864 ggcgcataagacagaagatcccgagccttaaaagagctgtgttagcaacctagaagaatcc 1923  
 |||||  
 Db 1501 ggcgcataagacagaagatcccgagccttaaaagagctgtgttagcaacctagaagaatcc 1560  
 QY 1924 ctgacttcttctgtgaagttgttctcaacgaacccggaagggatgtatgctgtgttcta 1983  
 |||||  
 Db 1561 ctgacttcttctgtgaagttgttctcaacgaacccggaagggatgtatgctgtgttcta 1620  
 QY 1984 agagaagctggtgtatgtgacagccttaaaagagaatgtgtcttcatgtatgataccta 2043  
 |||||  
 Db 1621 agagaagctggtgtatgtgacagccttaaaagagaatgtgtcttcatgtatgataccta 1680  
 QY 2044 ggaagcatalagagactccatgaacaaagcttagaanaaagttagagaagcgttgaaggaa 2103  
 |||||  
 Db 1681 ggaagcatalagagactccatgaacaaagcttagaanaaagttagagaagcgttgaaggaa 1740  
 QY 2104 agagaagctgtgacagaggtgtgttgaaggaatgttcaacaagcttccctgtgataccccc 2163  
 |||||  
 Db 1741 agagaagctgtgacagaggtgtgttgaaggaatgttcaacaagcttccctgtgataccccc 1800  
 QY 2164 ctgcttctgtcctcgtacgggtgcccctagtagctcgtctcgttacttaacgttgggct 2223  
 |||||  
 Db 1801 ctgcttctgtcctcgtacgggtgcccctagtagctcgtctcgttacttaacgttgggct 1860  
 QY 2224 tgccttaataataggttctgtccttctgttgaagaaacgggtgaggtccagtccaagtatg 2283  
 |||||  
 Db 1861 tgccttaataataggttctgtccttctgttgaagaaacgggtgaggtccagtccaagtatg 1920  
 QY 2284 gtacttaggaacagtagcaagagccttctgaacaaaggaagaaactgaacctcag 2337  
 |||||  
 Db 1921 gtacttaggaacagtagcaagagccttctgaacaaaggaagaaactgaacctcag 1974

RESULT 5  
 AAF7726  
 AAF7726 standard; DNA; 7333 BP.

XX AAF77726;  
 AC 23-MAY-2001 (first entry)  
 XX  
 DT Defective retroviral genome isolated from PK-15 cell line.  
 XX  
 DE Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;  
 KW ss.  
 XX  
 RW Unidentified.  
 XX  
 OS US6190861-B1.  
 XX  
 PN 20-FEB-2001.  
 XX  
 PD 13-DEC-1996; 96US-0766528.  
 XX  
 PF 14-DEC-1995; 95US-0572645.  
 XX  
 PR (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PA Fishman JA;  
 XX  
 PI WPI: 2001-256211/26.  
 DR P-PSDB: AAB73282, AAB73283, AAB73284.  
 XX  
 PT Assessing risk of endogenous retroviruses in clinical practice and in  
 PT xenotransplantation, complises using probe sequences derived from swine  
 PT or miniature swine retroviral genome -  
 XX  
 PS Claim 1; Fig 2; 127pp; English.  
 XX  
 CC The present invention relates to a method for screening a cell or tissue  
 CC for the presence or expression of a retrovirus (RV), comprising  
 CC contacting a target nucleic acid from the cell or tissue with a second  
 CC nucleic acid from the present invention (e.g. the present sequence or a  
 CC fragment thereof). The method is useful for RV detection and to assess  
 CC graft transplantation risk. Screening of animals allows the elimination  
 CC of donors with active replication of known viruses. Inactive proviruses  
 CC can be detected and inactivated, allowing identification and elimination  
 CC of potential human pathogens derived from swine in a manner not possible  
 CC in the outbred human organ donor population and is important to the  
 CC development of human xenotransplantation.  
 CC  
 CC Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;  
 XX  
 QY Query Match 63.7%; Score 1882.8; DB 22; Length 7333;  
 QY Best Local Similarity 80.1%; Pred. No. 0;  
 QY Matches 2409; Conservative 0; Mismatches 402; Indels 195; Gaps 9;

DB 4759 ctatctcgtacttttgaccacacacggtctgtgaagcgaag----- 4803  
 QY 361 tcatgtatcccaagtaagctggtccacctccgactccggtgtgagagccgaaaga 420  
 DB 4804 ----- 4803  
 QY 421 ctgagaatcccttaagcttcgctccatcgctgtgtcttaacttaacaactccc 480  
 DB 4804 -----gaatcccttaagcttcgctccatcggtgtgtcttaacttaacaactccc 4859  
 QY 481 caggccaagtaagcaagccttaagacagctcgaaccccatagaccttaacctacc 540  
 DB 4860 caagtaatgtgtaacgctctgtgacagccgaactcccatataaccttaactctccac 4919  
 QY 541 tggctaatattatgacctgtatagcgtgtcaactgttaataagcaactgaggtgtctct 600  
 DB 4920 tggtaacttaactgtactcgttaccaggtatataataaacaagcactcaagggtggtccc 4979  
 QY 601 agagggcaacctgtggtcgtacgtcatcttcgctccgattgttaacccgctgtta-- 658  
 DB 4980 ttggggacacctgtggtcgtatataatgtctccttcgactcaagtaacctgttccaat 5039  
 QY 659 -----aaagcaacctcccaacctagtcgtaagttatgtgtctgtccaggt---c 711  
 DB 5040 gaccaggaacacaccccgatgtaactcgtgtactcaggtgttttaagttgtccaggaacc 5099  
 QY 712 acggaagaagaagaataacgtgtggtgtcttgggaatccctctgtgaagatgtgagctgc 771  
 DB 5100 ccaataatgaagaatattgtgaaatccctcaggaatttcttgcagaagaatgtgagctgc 5159  
 QY 772 gtacctccaagatgtgagatgtgaatgtgcaatctctccaggaacgggttaaatc 831  
 DB 5160 atactctcaatgtatgtggaatttggaaatgtgcagttccagcaagaagatgaagtac 5219  
 QY 832 tctctgt-----caa 842  
 DB 5220 tctttgttaacaatccctaccagttataatcaattatgtgcatgtggtgagatgtgaa 5279  
 QY 843 ttccggtccgggcaagtaacaaatgtgaactataaagaatagatgtctccaccatc 902  
 DB 5280 gattgtcaacagcgtgttaaaaaagattg-taagaataagcaataagctgtcatcgt 5338  
 QY 903 agacttagatttctaaagaataagtttcaacttaaaaaaggaagaacggaataattcaaa 962  
 DB 5339 agacttagatttctaaagaataagtttcaacttaaaaaaggaagaacggaataattcaaa 5398  
 QY 963 gtgataaattgtatagctgtggaatagtttta--ataatgtggtggtgagcag 1019  
 DB 5399 gtggtgaaattgtatatctgtggtgaatagtgtaactatgtgaggtctctgtggaagaagaag 5458  
 QY 1020 gtccaattcaacattcgtccttaagatagagaaggggagacacccctgtgtgcaatgtg 1079  
 DB 5459 attgttctcgaattatcgtccttaagatagagaatcagatgtgaactccgtgtgtcatag 5518  
 QY 1080 acccgataaagtaactgtggtcgaacaggggtcccggtcgtgtggtgacacgttaacttgc 1139  
 DB 5519 accaataaaggttgtggtcgaacaggtacccccaatccaaga----- 5560  
 QY 1140 ggtgcccccaataaactcgtgtggtgtgacataacacggtccttaagcaacagtaccac 1199  
 DB 5561 -----acagaggtcatalctccttaacccctctgttatacaatacaactc 5602  
 QY 1200 tggattgtatcttaccacaacagcctagaactccccaagtgltctgttgaagacagaca 1259  
 DB 5603 tggattgtatcttaccacaacagcctagaactccccaagtgltctgttgaagacagaca 1259  
 QY 1260 ggaactctcaagtcctcaccagggagcttccaagcaatacaactcaccagccgctgagtc 1319  
 DB 5648 gaaacttttttagccctacccagggagcttccaagctcttaactccaagactccaagagtc 5707  
 QY 1320 caactctctgtgtgtgtgtgtatcctcaggggtcctctattatgaagggatgtgctaa 1379



FT		/tag=	g
FT		/note=	"putative ENV coding region (partial) as described in the specification"
FT	mat_peptide	4752..6722	
FT		/*tag=	h
FT		/note=	"ENV protein (partial)"
PX			
PN	W09721836-A1.		
PD			
XX	19-JUN-1997.		
PF			
XX	13-DEC-1996;	96MO-USJ9680.	
PR			
XX	14-DEC-1995;	95US-0572645.	
PA	(GEHO ) GEN HOSPITAL CORP.		
PI	Fishman JA;		
XX			
DR	WPI: 1997-3332804/30.		
PT	P-PSDB: AAM32091-W32095.		
PT	New nucleic acid from porcine retroviruses - used for detecting viruses in transplant or other tissue and for assessing risk of transmitting infection to graft recipient		
XX			
PS	Claim 16; Fig 2; 128pp; English.		
CC	This cDNA sequence represents a defective purified swine retrovirus found in PK-15 cells containing the putative coding regions for viral GAD, POL and ENV proteins. There are a few in frame stop codons and CC apparent frame shifts in the given coding sequence which alter features of the translation. This sequence and PCR fragments generated from the CC sequence (see AMF74812-T74882) could be used to screen organs for the CC presence of porcine retroviruses prior to xenotransplantation. CC Transplantation can increase the likelihood of retroviral activation if CC intact and infectious proviruses are present. The porcine retroviral CC sequence can be used to generate probes to determine the level (e.g. copy number) of intact (i.e. potentially replicating) porcine provirus CC sequences in a strain of xenograft transplantation donors. It can be CC used to detect mutations, genetic lesions or viral recombinants and to CC determine the histological localisation of activated retroviruses. Using CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear CC cells, infectivity titration and susceptibility testing can be CC performed. Ultimately animal donors without intact porcine retroviral CC sequences or a lower copy number of viral elements could be selected. CC XX		
SO	Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other;		
	Query Match	61.2%; Score 1809.6; DB 18; Length 7393;	
	Best Local Similarity	78.5%; Pred. No. 0;	
	Matches 2407; Conservative	0; Mismatches 404; Indels 255; Gaps 10;	
OY	1 tgccttttaaggttgagaacacccttggaaagtgtgggtgaaccccatgatgtcct 60		
DB	4462 tgccttttaagggtgaggaacaccttgacaagtlttgggtcgccccataatgtcct 4521		
OY	61 acggggagacccccccggtgtgtagaatgtcttcgtacatatgttgtatgtgcgtctt 120		
DB	4522 acggggagacccccccggtgtgtagaatgtcctttgcaatatgttgtatgtgcgtctt 4581		
OY	121 ccagacttgttctctataggtctcaaggcgctcgtatgtgtgtgagcaagcgttgaagc 180		
DB	4582 ccagacttgttctctataggtctcaaggcgctcgtatgtgtgtgagcaagcgttgaagc 4641		
OY	181 agctcggagagcctaactaagagagagagacttgcgaattcacatcgtccccaagtgt 240		
DB	4642 agctcggagagcctaactaagagagagacttgcgaattcacatcgtccccaagtgt 4698		
OY	241 gagattcagctatgtttaagcgcacacgttgaagaaacctcgagactcgttgaagagac 300		
DB	4699 gagattcagctatgtttaagcgcacacgttgaagaaacctcgagactcgttgaagagac 4758		

QY	1320	caattcttctgttggttcgttctatcctcaaggctcctattatlgaaaggatgctaa	1379
Db	5708	taacctctctgttggtcttgcttgaagcttgcggccacctactatlgaaaggatgtctag	5767
QY	1380	agaagaataatcgaatgtgacaaaggagcctagaaatacaatactaatgtgggttccggaa	1439
Db	5768	aggagggaaattcaattgacaaaggaaacataagacccaatgtaactgtggatcccaaa	5827
QY	1440	taagcttaacctcactgaatlttcggggaaagggaatgcagatagaaagaagcccccat	1499
Db	5828	taagcttaaccttaactgaagtttctotgaaaaaggcacctgcataggatgtgtcccatc	5887
QY	1500	ccaacaaaccttggctatagtactgtgtttatgtgcaggcctcaagaatacagatatt	1559
Db	5888	ccaacaaaccttggtaacacacactgaagcctttaaactgaacctgtgagtgacatattc	5947
QY	1560	agtaacctgttatacaagcgtggtgggcacatgaactgtgttaacccctgtttccac	1619
Db	5948	ggtactcgtgttatgacagcgtggtgggcacgttaactcgtgattaacccctgtgtccac	6007
QY	1620	ctcagcttctaaacatcacaagaattgtgtgtcattgtgcacaaatcgtccccggagtta	1679
Db	6008	ctgtgttctcaaccaacataagaactttgtgtgtatgtgtccaaattgtccccgggtga	6067
QY	1680	ctaccatcctggaaggaaagtgtctcttgatgaatatgactatcgtgtataccgcaaaaag	1739
Db	6068	cttactatcccgaaaaaagcagctcttgatgaatatgactatgaatactgcgcaaaaag	6127
QY	1740	agaacccgtatccttaaccttaactgaatgtgcgattagaagcgcgtgttcgctag	1799
Db	6128	agagcccatatccctgaacactagctgttaattgtctcgatgtggagtgctgcgaagcgtg	6187
QY	1800	aacagggacaagctgtcccttgatcaacagagcacagacgtatagaagaagactgtgtgact	1859
Db	6188	aacagggacaagctgtccctatcatcacagacacagacgtatagaagaagactgtgtaact	6247
QY	1860	acatgcgcgcatgaagaagaatctccgagccttaagaagatctgttgcacctagaaga	1919
Db	6248	acatcgaaattgtacaaggaaatctccaagccttaagaataactgttcaagtaacctgga	6307
QY	1920	gtcccttactctcttgcgaagtgtgttacaagaacggggagatlaagctcgcgtt	1979
Db	6308	atcccttaacctcttatctcgaagtgttcttaagaacacgaagggggtttagctcgttatt	6367
QY	1980	tctaagaagaagtgtggttatgtgcagccttaaaagaagaatgtgtctctatgtagaaca	2039
Db	6368	tctaagaagaagggtgtatgtgtagccttaaaagaagaatgtgtctctatgtagaaca	6427
QY	2040	ctcagggcgcatagaagatcatatgaacaagaattagaaaaaagttagaagcgctgaag	2099
Db	6428	ctcagggcgcatatagaagatcatatgaacaagaattagaaaaaagttagaagcgctgaag	6487
QY	2100	ggaagaagagcgtcgaacaaagggtgttgaagaagatgttcaacaggcctccttgatgac	2159
Db	6488	ggaagaagagcgtcgaacaaagggtgttgaagaagatgttcaacacaggcctccttgatgac	6547
QY	2160	caaccctgttctgtctgcgcggggccctcctatagtcctgtcctgttacttiaaagtltg	2219
Db	6548	caaccctgttctgtcctctacgcggggccctcctatagtcctgtcctgttacttiaaagtltg	6607
QY	2220	gctctgcttaatatagttgtgtgctctgtttagagaacagatgtgagtgacgtccaagat	2279
Db	6608	gctctgcttaatatlaaagttgtgtgctctgtttagagaacagatgtgagtgacgtccaagat	6667
QY	2280	catgctactaagcacaagtatcaagagcctctgagccaaggagaaactgaaaccttaagcc	2339
Db	6668	catgctactaagcacaagtatcaagagcctctgagccaaggagaaactgaaaccttaagcc	6727
QY	2340	ttcccaagttcctaagaattagaactatlaacaagaacaagaatgtggagatgaagaatgaaa	2399
Db	6728	ttcccaagttcctaagaattagaactatlaacaagaacaagaatgtggagatgaagaatgaaa	6787
QY	2400	atgcaaacctaacacctccagaaccccgaggatgaataaaaaagctcctaattgtccccgaat	2459

Db	6788	atgcaacttaacccctccacagaccacaggaattaaataaaagctctaattgcctccgat	6847
QY	2460	tacgaacctgtgtgctgtgccagtaaataggtagaagtgacacattcctattgttccagg	2519
Db	6848	tccagaccctgtgctgtgccagtaaataggtagaagtgacacattcctattgttccagg	6907
QY	2520	cctgtctatcttcggccctaagtaagaataacaggaataggttgactatctgtc-----	2570
Db	6908	cctgtctatcttcggccctaagtaagaataacaggaataggttgactatctgtc-----	6967
QY	2571	-----tactgtgat	2579
Db	6968	tctgttaaacctgactggaaccatagaagaattgattacaacattgcagccctatctgat	7027
QY	2580	tctgttaaacctgactggaaccatagaagaattgattacaacattgcagccctatctgat	2639
Db	7028	tctgttaaacctgactggaaccatagaagaattgattacaacattgcagccctatctgat	7087
QY	2640	atctcaactgcaactcttcaactctgcgccagagagccacagacagatgaggactccgagct	2699
Db	7088	atctcaactgcaactcttcaactctgcgccagagagccacagacagatgaggactccgagct	7147
QY	2700	attttaaaatgattgtgtccacagagacgctgctctcgatatatttaaaatgattgtccat	2759
Db	7148	attttaaaatgattgtgtccacagagacgctgctctcgatatatttaaaatgattgtccat	7207
QY	2760	ggaagcgaggctctcgatatatttaaaatgattgtgttgagacacagagctttgtga	2819
Db	7208	ggaagcgaggctctcgatatatttaaaatgattgtgttgagacacagagctttgtga	7267
QY	2820	accocataaaagctgttcccgatctccgacttcggggccgcagtcctctacacccttgcgtg	2879
Db	7268	accocataaaagctgttcccgatctccgacttcggggccgcagtcctctacacccttgcgtg	7327
QY	2880	gtacgactgtggggcccccagagcgcttggaataaaatcctcttgcgtttggatcaaaaa	2939
Db	7328	gtacgactgtggggcccccagagcgcttggaataaaatcctcttgcgtttggatcaaaaa	7387
QY	2940	aaaaaa2945	
Db	7388	aaaaaa7393	
RESULT 7			
AAV82748			
ID	AAV82748	standard; DNA; 2462 bp.	
AC	AAV82748;		
XX			
DT	25-FEB-1999	(first entry)	
XX			
DE	Pig endogenous retrovirus (PERV)-A envelope (env) gene region.		
XX			
KW	Pig endogenous retrovirus; PERV-A: envelope protein; PERV-B: subtype;		
KW	probe; primer: detection; retrovirus; human tissue; xenotransplant;		
KW	primary porcine tissue; human cell line; porcine cell line; ss.		
XX			
OS	Pig endogenous retrovirus.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	211..2193	
FT		/tag= a	
FT		/product= envelope_protein	
XX	W09853104-A2.		
XX			
PD	26-NOV-1998.		
XX			
PF	18-MAY-1998.	98WO-G801428.	
XX			
XX	16-MAY-1997;	97GB-0010154.	
XX			

PA (MED-1) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI: 1999-045324/04.

DR P-PSDB; AAW85453.

XX Newly isolated nucleic acid probe capable of hybridizing to either  
PT the PERV-A or PERV-B env gene - useful in the detection of  
PT retroviruses, and their subtypes, in a sample of porcine/human  
PT tissue

XX Claim 6; page 20-21; 36pp; English.

XX The present sequence encodes a P1g endogenous retrovirus (PERV)-A  
CC envelope protein. PERV exists in two different subtypes, PERV-A and  
CC PERV-B. The differences are reflected in sequence divergence in the  
CC envelope genes. Probes and primers can be derived from the envelope  
CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
CC a method to detect retroviruses in a sample of porcine/human tissue,  
CC particularly primary porcine tissue and human cell lines that have been  
CC cultivated in the presence of a porcine cell line, or human tissue from  
CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
CC one of the PERV env genes can also be determined.

XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

Query Match 52.9%; Score 1563.4; DB 20; Length 2462;

Best Local Similarity 79.5%; Pred. No. 0;

Matches 2000; Conservative 0; Mismatches 401; Indels 116; Gaps 8;

QY 151 tcgagtggtgaggaagcaagagctggaagcagctccggagagcctcctcagagagaag 210  
DB 1 tcgagtggtgaggaagcaagagctggaagcagctccggagagcctcctcagagagaag 57  
QY 211 actgcaagttccacatcgctcccaagttgagatcagctatgtttagagccgacgag 270  
DB 58 actgcaagttccacatcgctcccaagttgagatcagctatgtttagagccgacgag 117  
QY 271 caggaacactcgagactcgtggaagagacattatcgtactttgacacacaaag 330  
DB 118 caggaacactcgagactcgtggaagagacattatcgtactttgacacacaaag 177  
QY 331 ctgtgaaagtcgaagatctccactgagatccatgacatccacgttaagctggcgac 390  
DB 178 ctgtgaaagtcgaagatctccactgagatccatgacatccacgttaagctggcgac 237  
QY 391 ctccgagactcggtgagagccgaagaagactggaatcccttaagctcgctccatc 450  
DB 238 ctccgagactcggtgagagccgaagaagactggaatcccttaagctcgctccatc 297  
QY 451 gctcgttcttacttaacaataactcccaagccagtagtaaacgcttataagacagc 510  
DB 298 gctcgttcttacttaacaataactcccaagccagtagtaaacgcttataagacagc 357  
QY 511 tcgagacccatagactcttataccctactgctgtagtatttgaacctgtgaagagtc 570  
DB 358 cgcgaatcccatagactcttataccctactgctgtagtatttgaacctgtgaagagtc 417  
QY 571 acttaataagactcagagtggtgctctcctagagagcactggtgagcttgaacttcc 630  
DB 418 aataataagactcagagtggtgctctcctagagagcactggtgagcttgaacttcc 477  
QY 631 tgcctcgattgaatacccgctgtta-----aaagcacactcccaactagtcgt 684  
DB 478 tgcctcgattgaatacccgctgttcaatgacagccacaccccgagtgtaactcgt 537  
QY 685 agtatgggtcttatctgctgcagag---cacagagaagaagaaactgtgggggtct 741  
DB 538 gcttaaggggttttaacgtttgcccagagaccccaataatgaagaatatttggagaactc 597  
QY 742 gggaatcctcttctgtagagatgagactcgtaacctccaacgatagagactggaatg 801

DB 598 caggaattctcttctgcaagcaatgagcgtgacatcttaactgtagtgaattggaatg 657  
QY 802 ccgactctctccagagccggtataaactctccttctgaactc----- 846  
DB 658 ccgactctctccagagccggtataaactctccttctgaactc----- 846  
QY 847 -----ggccggggaagtaacaatgatga 872  
DB 718 caatttaattgacctgagatgagatggaagatggaacagcggtgtcaaaaagatg 776  
QY 873 actataaagataagagctctcccaactgagacttagatctataaagataaatttccac 932  
DB 777 acgaataaagataaagatgagatgagatggaagatggaacagcggtgtcaaaaagatg 836  
QY 933 tgaanaaggaacacaggaatataatcaaatgtagtaaatgtagtgcgtgggaatg 992  
DB 837 tgaanaaggaacacaggaatataatcaaatgtagtaaatgtagtgcgtgggaatg 896  
QY 993 tttta-----ataatgctgggggagaggtgacacttaacactgctgagataga 1049  
DB 897 gtaactatgagagctctgggagagaagaagagatcttctgactatctgcccagataga 956  
QY 1050 gacgggacagagacccctgtgcaatgagacccataaagtaactgctgagagggcc 1109  
DB 957 aactcagatgagacccctgtgcaatgagacccataaagtaactgctgagagggcc 1016  
QY 1110 cccgacctgagagccacgcaataactgctgagacccataaactgctgagggcc 1169  
DB 1017 tccaatccagagac-----gagggccac 1040  
QY 1170 cataacacagccgcttagacacagtaacacagtaactgtagtcccaacacagcctagaa 1229  
DB 1041 tccataacccctttagtataaataacacactctgtagtcccaacacagcctagaa 1088  
QY 1230 ctcccaagtgcttctgttgaagacagagacagactcttcaagctccacagagagctt 1289  
DB 1089 ---gctaacatcacatataaacaagggcgagaaactttagcctcaccaagagagctt 1145  
QY 1290 ccaagcactcaactcagacccgagacccctgagaccccttctgttggcttgcatactc 1349  
DB 1146 tcaagctcttaactccac 1205  
QY 1350 agggcctcttataag 1409  
DB 1206 gggccacacttaactag 1265  
QY 1410 tagaatacagtagtactgag 1469  
DB 1266 tagaatacagtagtactgag 1325  
QY 1470 gggagacatag 1529  
DB 1326 agggacactgag 1385  
QY 1530 ttagagcagagcctcag 1589  
DB 1386 cttaatcgacactctgag 1445  
QY 1590 caatctgag 1649  
DB 1446 taactatgag 1505  
QY 1650 tgcctgag 1709  
DB 1506 cgttatgag 1565  
QY 1710 atatactatcggtataacag 1769  
DB 1566 atatactatcggtataacag 1625  
QY 1770 gctcagatag 1829

Db 1626 gctcgatgtggtgctgacgagcgtggaacaggaacgctgccttaacacagagacc 1685  
 QY 1830 acagacgtctagagaaagactgtgtgacatcagtcgacatgacagaagaatccagac 1889  
 Db 1686 gaacagctggaagaaagacttagtaacctacatcgaattgtaacgaagatcccaagc 1745  
 QY 1890 cttaagagatctgttaacaccttagaagatccctgactcttctgttgaagtgcttct 1949  
 Db 1746 cctagaaaaactgtcagtaacctggaagaaatccctacccctctatctgaagtctct 1805  
 QY 1950 acagaacggagggatagatagatctgctgttcttaagaagaagtggttatctgacgctt 2009  
 Db 1806 acagacagaaggggttagatctgttattcttaagaagaaggggttatgtgagctt 1865  
 QY 2010 aaagaagaagatgttctctatgtatgacatcagagagacacacacactcatalgaa 2069  
 Db 1866 aaagaagaagatgttctctatgtatgacatcagagagacacacacactcatalgaa 1925  
 QY 2070 gcttagaaaaagttagagagcgctgacaggaaggaagagagcgctgacaggggtgttga 2129  
 Db 1926 gcttagaaaaagttagagagcgctgacaggaaggaagagcgctgacaggggtgttga 1985  
 QY 2130 aggaatgttcaacagagctccttgatgacacacacacacacacacacacacacacac 2189  
 Db 1986 aggaatgttcaacagagctccttgatgacacacacacacacacacacacacacacac 2045  
 QY 2190 agtatgctgtcctgttacttaacatggtgagcctgtcttaataatgattgttgcctt 2249  
 Db 2046 agtatgctgtcctgttacttaacatggtgagcctgtcttaataatgattgttgcctt 2105  
 QY 2250 tgttagagaagagatgagtgacagcagatcagatcagatcagatcagatcagatcagat 2309  
 Db 2106 tgttagagaagagatgagtgacagcagatcagatcagatcagatcagatcagatcagat 2165  
 QY 2310 tcttagagcaagggagaaactgacatcagccttccagcttctaagatagatcactaa 2369  
 Db 2166 tcttagagcaagggagaaactgacatcagccttccagcttctaagatagatcactaa 2225  
 QY 2370 agacaagaatggtggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 2429  
 Db 2226 agacaagaatggtggaatgaaagatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 2285  
 QY 2430 gtttaataaagaacttaaaatgcccccaatcagacacacacacacacacacacacacac 2489  
 Db 2286 gtttaataaagaacttaaaatgcccccaatcagacacacacacacacacacacacacac 2345  
 QY 2490 tagaaggtcacaacttctatgttccagggcgtgtatcctgtggtcctaagtaagatacag 2549  
 Db 2346 tagaaggtcacaacttctatgttccagggcgtgtatcctgtggtcctaagtaagatacag 2405  
 QY 2550 gaagtgtgacttaactgcttctggtatctggttaactgactgacacataaga 2606  
 Db 2406 gaagtgtgacttaactgcttctggtatctggttaactgactgacacataaga 2462

RESULT 8  
 AAC67022  
 ID AAC67022 standard; DNA; 4918 BP.  
 XX  
 AC AAC67022;  
 XX  
 DT 27-MAR-2001 (first entry)  
 XX  
 DE PERV env protein coding sequence SEQ ID NO: 22.  
 XX  
 KW Xenotransplantation; infectious agent; vaccine; ds.  
 XX  
 OS Porcine endogenous retrovirus.  
 XX  
 PN MO200071726-A1.  
 XX  
 PD 30-NOV-2000.  
 XX

PF 24-MAY-2000; 2000MO-US14296.  
 XX  
 PR 24-MAY-1999; 99US-0135631.  
 XX  
 PA (MAYO-) MAYO MEDICAL VENTURES.  
 XX  
 PI Federspiel MJ;  
 XX  
 DR WPI; 2001-032041/04.  
 XX  
 PT Inhibiting or preventing infectious agent transmission in mammalian  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 PS Claim 16; Page 109-111; 144pp; English.  
 CC  
 CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 CC  
 XX  
 SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

Query Match 52.5%; Score 1552.2; DB 22; Length 4918;  
 Best local similarity 73.2%; Pred. No. 0;  
 Matches 2257; Conservative 0; Mismatches 603; Indels 223; Gaps 12;

QY 1 tgcctttaggggtltaggaacacccctgagacagttgggtgacccctatgaatgtcct 60  
 Db 1824 tgcctttaggggtltaggaacacccctgagacagttgggtgacccctatgaatgtcct 1883  
 QY 61 acgggggagaccccccccggtgtagaagaattgtctctgataatgctgtatgtcgtctt 120  
 Db 1884 acgggggagaccccccccggtgtagaagaattgtctctgataatgctgtatgtcgtctt 1943  
 QY 121 cccagccttctgtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180  
 Db 1944 cccagccttctgtctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2003  
 QY 181 agctccggagagcctcactcagagagagagagactgtgcaatccatcgtctccaagt 240  
 Db 2004 agctccggagagcctcactcagagagagagagactgtgcaatccatcgtctccaagt 2063  
 QY 241 gagatcagctcagtttagagaccccgctggaagaactggaactgtgtggaagagac 300  
 Db 2064 gagatcagctcagtttagagaccccgctggaagaactggaactgtgtggaagagac 2123  
 QY 301 ctatctcgtactcttgac 360  
 Db 2124 ctatctcgtactcttgac 2183  
 QY 361 tccatgcatcccaagtttagcgtgacacacacacacacacacacacacacacacacacac 420  
 Db 2184 tccatgcatcccaagtttagcgtgacacacacacacacacacacacacacacacacacac 2243  
 QY 421 ctgagaatcccttaagattcgctcgaacgctggttcttactcttaacataacacacac 480  
 Db 2244 ctgagaatcccttaagattcgctcgaacgctggttcttactcttaacataacacacac 2303  
 QY 481 caagcagtagtaaacgcttatagacagcctggaaccccaataacacacacacacacacac 540  
 Db 2304 caagcagtagtaaacgcttgtagaagcccggaaccccaataacacacacacacacacac 2363  
 QY 541 tgcctgattatgac 600  
 Db 2364 tgcctgattatgac 2423  
 QY 601 aagagcactgtgtgac 658  
 Db 2424 ttgggagacgtgtgac 2483





Db 4581 tgcgcacatgtgtactccttaacaaatlaattatgtgtccgcgaagcgcgccctc 4640  
 QY 2736 gatatataaagtattgttcacatgagcgcggtctcgcataltttaaatgatgtgt 2795  
 Db 4641 gaagtttaattgactcgt-----ttgtatattttgaatgatgtgt 4685  
 QY 2796 tgtacgcacagcgtttgtgttgacacccataaagctgtcccgatccgcactcggtgc 2855  
 Db 4686 tgltaagcgcggtctgtgttgacacccataaagctgtcccgatccgcactcggtgc 4745  
 QY 2856 cgcagctcctacccctcgtgtgtgtacgaactgtggtcccgacggtcggtgataaata 2915  
 Db 4746 cgcagctcctacccctcgtgtgtgtacgaactgtggtcccgacggtcggtgataaata 4805  
 QY 2916 tccctctgtcgtttgcatacaaa 2938  
 Db 4806 tccctctgtcgtttgcatacaaga 4828

## RESULT 9

AA77727 standard; DNA; 8132 BP.

AA77727;

23-MAY-2001 (first entry)

Nucleotide sequence of a retrovirus found in miniature swine.

Retrovirus; graft transplantation; xenotransplantation; miniature swine;

ss.

Unidentified.

US6190861-B1.

20-FEB-2001.

13-DEC-1996; 9605-0766528.

14-DEC-1995; 9505-0572645.

(GEHO) GEN HOSPITAL CORP.

Fishman JA;

WPI; 2001-256211/26.

P-PSDB; AAB73285, AAB73286, AAB73287.

Assessing risk of endogenous retroviruses in clinical practice and in

xenotransplantation, comprises using probe sequences derived from swine

or miniature swine retroviral genome

Claim 1; Fig 3; 127pp; English.

The present invention relates to a method for screening a cell or tissue

for the presence or expression of a retrovirus (RV), comprising

contacting a target nucleic acid from the cell or tissue with a second

nucleic acid from the present invention (e.g. the present sequence or a

fragment thereof). The method is useful for RV detection and to assess

graft transplantation risk. Screening of animals allows the elimination

of donors with active replication of known viruses. Inactive proviruses

can be detected and inactivated, allowing identification and elimination

of potential human pathogens derived from swine in a manner not possible

in the outbred human organ donor population and is important to the

development of human xenotransplantation.

Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

Query Match 50.4%; Score 1489.2; DB 22; Length 8132;

Best Local Similarity 72.2%; Pred. No. 0;

Matches 2140; Conservative 0; Mismatches 713; Indels 113; Gaps 11;  
 QY 1 tgcatttagggtagaacacccctgagacagtttgggtgacccctatgatatgtcct 60  
 Db 5257 tgcatttagggtagaacacccctgagacagtttgggtgacccctatgatatgtcct 5316  
 QY 61 acgggggagaccccccgtgtgttagaaattgtctctgtacatagtcgtatgtgtcctt 120  
 Db 5317 acgggggagaccccccgtgtgttagaaattgtctctgtacatagtcgtatgtgtcctt 5376  
 QY 121 cccagccttgtctctcttagcacaagcgctcgaatgtgtgtgtggaacagcgtagaagc 180  
 Db 5377 cccagccttgtctctcttagcacaagcgctcgaatgtgtgtgtggaacagcgtagaagc 5436  
 QY 181 agctccggagagcctactcagagagagagactgtcaaatctccacatcgtctccaagtgt 240  
 Db 5437 aactccggagagcctactcagagagagagactgtcaaatctccacatcgtctccaagtgt 5496  
 QY 241 gagatcagatctatgttagacgcacacgcgtgtaaaagtcgaagaaatccacactga 300  
 Db 5497 gagatcagatctatgttagacgcacacgcgtgtaaaagtcgaagaaatccacactga 5556  
 QY 301 ctatctcgtactcttgacacacacacgcgtgtaaaagtcgaagaaatccacactga 360  
 Db 5557 ctatctcgtactcttgacacacacacgcgtgtaaaagtcgaagaaatccacactga 5616  
 QY 361 tccatgcatcccaagtgtaagctgtgcgcacactccgcactcgggtgtgagagcggaagaa 420  
 Db 5617 tccatgcatcccaagtgtaagctgtgcgcacactccgcactcgggtgtgagagcggaagaa 5676  
 QY 421 ctgagatcccttaagctgtgcctcgcacactcgtgtcttacttacttaacataactccc 480  
 Db 5677 ctgagatcccttaagctgtgcctcgcacactcgtgtcttacttacttacttaacataactccc 5736  
 QY 481 cagcgcaagttagtaaacgcctttagacagctgcacacccccaatagacacttaccctacc 540  
 Db 5737 cagcgcaagttagtaaacgcctttagacagctgcacacccccaatagacacttaccctacc 5796  
 QY 541 tggcgtatattgacccctgatacaggtgtgtaactgttaaatagcaactcgaagtggtgtcct 600  
 Db 5797 tggcgtatattgacccctgatacaggtgtgtaactgttaaatagcaactcgaagtggtgtcct 5856  
 QY 601 agagcacctgtgtgacccctgatacaggtgtgtaactgttaaatagcaactcgaagtggtgtcct 660  
 Db 5857 agagcacctgtgtgacccctgatacaggtgtgtaactgttaaatagcaactcgaagtggtgtcct 5913  
 QY 661 agcacactcccaactagtcgt 717  
 Db 5914 agcacactcccaactagtcgt 717  
 QY 718 aagaggaataactgtggt 777  
 Db 5974 aagaggaataactgtggt 6033  
 QY 778 tccaacagatgagacactggaatgagcgcgtctctcctccagagacgggtlaaattctcctt 837  
 Db 6034 tccaacagatgagacactggaatgagcgcgtctctcctccagagacgggtlaaattctcctt 6093  
 QY 838 gtcaa-----ttccggccgggcaagtaacaaatgataagaaactataaagat 885  
 Db 6094 gtcaa-----ttccggccgggcaagtaacaaatgataagaaactataaagat 885  
 QY 886 aagagctctcccccagactgagactgagactgagactgagactgagactgagactgagactgag 945  
 Db 6154 aagagctctcccccagactgagactgagactgagactgagactgagactgagactgagactgag 6213  
 QY 946 caggaataatccaagaatgataaagtgtatgtgactgtggaatagattttta--ataa 1002  
 Db 6214 caggaataatccaagaatgataaagtgtatgtgactgtggaatagattttta--ataa 1002  
 QY 1003 tatggcgggagagcaaggttccacttcaacttgcacttgcacttgcacttgcacttgcacttgcact 1062  
 Db 6274 tatggcgggagagcaaggttccacttcaacttgcacttgcacttgcacttgcacttgcacttgcact 6330

QY 1063 cccctctgtgcaatgaggaccggaataaagtactgtctgaacagggccccgcccgtgag 1122  
 Db 6331 cctcaatggtatagaccgaataagcttctgaacgggtcaaaagaccccaacc----- 6385  
 QY 1123 ccacgcataactgtgcggtgccccaaataactgcgtgcgcgtgacatacaacagccg 1182  
 Db 6386 -----aaagacacagagca 6399  
 QY 1183 cctgaacagctaacactgtgattgattcctaccacaacgcgtagaactcccaagtggt 1242  
 Db 6400 tcccttaacataacttctgtgatacgaaccacagctgagtctaacagca----- 6445  
 QY 1243 cctgttaagaacagacagagactctcagtcatacccaaggagcttcccaagcatcaac 1302  
 Db 6446 -cgactaaatgagggaacaaacttttagcctcaccagagagcttttcaagctctaac 6504  
 QY 1303 tccacgacccctgatacgaactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1362  
 Db 6505 tccacgacccctgatacgaactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 6564  
 QY 1363 tatagaggagatgtgttaagaagaataatcaatgtgacccaagaagatagaatcaatgt 1422  
 Db 6565 tatagaggagatgtgttaagaagaataatcaatgtgacccaagaagatagaatcaatgt 6624  
 QY 1423 acatgagggtgccggaataagcttaacccctgaactgaagttccgggaaggagatgata 1482  
 Db 6625 acatgagggtgccggaataagcttaacccctgaactgaagttccgggaaggagatgata 6684  
 QY 1483 ggaagaagctcccccaccccaacaacactgtgctatagtactgtgtgtgtgtgtgtgtgt 1542  
 Db 6685 ggaagaagctcccccaccccaacaacactgtgctatagtactgtgtgtgtgtgtgtgtgt 6744  
 QY 1543 tccgaataatcagattatgaactgtgtgtataacaggtgtgtgtgtgtgtgtgtgtgtgtgt 1602  
 Db 6745 tccgaataatcagattatgaactgtgtgtataacaggtgtgtgtgtgtgtgtgtgtgtgtgt 6804  
 QY 1603 accccctgtgttccacccctgaacttccaaaccccaacattgtgtgtgtgtgtgtgtgt 1662  
 Db 6805 accccctgtgttccacccctgaacttccaaaccccaacattgtgtgtgtgtgtgtgtgt 6864  
 QY 1663 atcgtcccccagtgacttaacacccctgaggaagtggtgtgtgtgtgtgtgtgtgtgtgt 1722  
 Db 6865 atcgtcccccagtgacttaacacccctgaggaagtggtgtgtgtgtgtgtgtgtgtgtgt 6924  
 QY 1723 tataccgacccaaaagagagaccccgatcccttaacccctgaggtgtgtgtgtgtgtgtgt 1782  
 Db 6925 aatcatcagacaaaagagagaccccgatcccttaacccctgaggtgtgtgtgtgtgtgtgt 6984  
 QY 1783 acggtcgt 1842  
 Db 6985 gtgtcagcaggt 7044  
 QY 1843 aaaggaactgt 1902  
 Db 7045 acaggaactgt 7104  
 QY 1903 gtttagcaacttaagaaggttccctgacttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1962  
 Db 7105 gtttagcaacttaagaaggttccctgacttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 7164  
 QY 1963 ggaatagatcgt 2022  
 Db 7165 ggt 7224  
 QY 2023 tgcctctatgtatgaactaactaagagccatcaagactcaatacaacagcttagaanaag 2082  
 Db 7225 tgcctctatgtatgaactaactaagagccatcaagactcaatacaacagcttagaanaag 7284  
 QY 2083 tttagaagagcgtcgaagaagagagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142  
 Db 7285 tttagaagagcgtcgaagaagagagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 7344

QY 2143 aggtcctctgtgatgaccacccctgtctctgtctgaacggggcccttagtctgtctc 2202  
 Db 7345 aggtcctctgt 7404  
 QY 2203 ctgttactaagttgt 2262  
 Db 7405 ctgttactaagttgt 7464  
 QY 2263 gttgagtgagctcagatcgt 2322  
 Db 7465 ataatgtcagctcagatcgt 7521  
 QY 2323 gaaactgaccttaagctcctccagttcctaagattagaactataaacaagacagaagtgtg 2382  
 Db 7522 gaaactgaccttaagctcctccagttcctaagattagaactataaacaagacagaagtgtg 7581  
 QY 2383 ggaatgaagaagatgaanaatgcaaccttaacccctccagaacccaagagttataaanaagc 2442  
 Db 7582 ggaatgaagaagatgaanaatgcaaccttaacccctccagaacccaagagttataaanaagc 7641  
 QY 2443 tctaatgccccggaattacaacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2502  
 Db 7642 cagagttgttccctcatalaggaanaagatgaattgtgtgtgtgtgtgtgtgtgtgtgtgt 7701  
 QY 2503 ttcctaatgttccagggcctgtctatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2561  
 Db 7702 taaataagcctctgt 7761  
 QY 2562 c-----taatcgtctatcgt 2617  
 Db 7762 caagaanaagaggttcttaacgt 7821  
 QY 2618 acattgacagccctgt 2677  
 Db 7822 atgttgaatgt 7881  
 QY 2678 gcaagtgcgagctcccgagctatcttaaatgattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2737  
 Db 7882 cgcgaatgt 7941  
 QY 2738 taatttaaaatgt 2797  
 Db 7942 agtttaaaatgt 7966  
 QY 2798 tgaacgacaggt 2857  
 Db 7987 taaagcgcgggt 8046  
 QY 2858 cagtccttaacccctgt 2917  
 Db 8047 cagtccttaacccctgt 8106  
 QY 2918 ctctgt 2943  
 Db 8107 ctctgt 8132

RESULT 10  
 AAC67023  
 ID AAC67023 standard; DNA: 7873 BP.  
 AC AAC67023;  
 AC 27-MAR-2001 (first entry)  
 DE PERV env protein coding sequence SEQ ID NO: 23.  
 XX Xenotransplantation; infectious agent; vaccine; ds.  
 OS Porcine endogenous retrovirus.  
 XX MO200071726-A1.



Db 6277 ctgttttccaacaaactaaagacttttcgtatgtgtccaaatgtccccggtgtta 6336  
 Qy 1680 ctaccatcttggaagatggttccttgatgaatatactacatcgttlaaccgacaagaag 1739  
 Db 6337 ctactatcccgaaaaagcagtccttgatgaatatactacatcgttlaaccgacaagaag 6336  
 Qy 1740 agaaccgtatcccttaccatcgttgaatgtcctcgtatgaagcagcgtttggtgtag 1799  
 Db 6397 agagcccatcccttgacacatcgttgaatgtcctcgtatgaagcagcgtttggtgtag 6456  
 Qy 1800 aacaggagcagctgcccgtacacagcagcagcagcagcagcagcagcagcagcagcagc 1859  
 Db 6457 aacaggagcagcgtcccttaacacagcagcagcagcagcagcagcagcagcagcagcagc 6516  
 Qy 1860 acatcgcgcacatgacagaagatctccgagccttaagagagcgttgaacacacagcagcagc 1919  
 Db 6517 acatcgaattgttaacaggaagatctcccaagccttaagaaatcgttcaatcccgagcagcagc 6576  
 Qy 1920 gtccctgactcttctgtctgaagtgttctacacagacgcgagggatagatctctgtgt 1979  
 Db 6577 atcccaaccctctatctgtgaagtgttctacacagacgcgaggggttagatctctgtat 6636  
 Qy 1980 tctaagaagaagtgtgtatgtgtcagccttaagaagaagaatgtgtctctatgtatga 2039  
 Db 6637 tctaagaagaagtgtgtatgtgtcagccttaagaagaagaatgtgtctctatgtatga 6696  
 Qy 2040 ctcaagagccatcagagacatcctcagcaacagccttagaanaa----- 2081  
 Db 6697 ttcaggggcacatcagagacatcctcagcaacagccttagaanaaagcctgagagcgcgcgcg 6756  
 Qy 2082 ----- 2081  
 Db 6757 ggtctcgacaacacagacaggttctgttcaatlaaagaactgtcgaagtaaccgt 6816  
 Qy 2082 -----gttagagagcgtcgcgaaggaagaagagcgtgaca 2117  
 Db 6817 gagctaaagccagcgttagagaagaaggtcgcgaaggaagaagaactacatca 6876  
 Qy 2118 ggggtgtgttgaagatgtgtcaacaggtcccttgatgacccacccctgtctgtctc 2177  
 Db 6877 agagtgtgttgaagatgtgtcaacaggtcccttgatgacccacccctgtctgtctc 6936  
 Qy 2178 gacggggcccttagttagtctgtcctgttactacagcttgggctgttaataatag 2237  
 Db 6937 aacaggaccccttaatagctccctcctgttactacagcttgggctgttaataatag 6996  
 Qy 2238 gttgtgtgcttgttgaagacaggtgagtgacgtccacatcgttacttaggaca 2297  
 Db 6997 gttatgtgcttctacatagagaacgaatgacgtccacatcgttacttaggaca 7056  
 Qy 2298 gtaccagagccttctgagcacaagaaactgacgtccacatcgttacttaggaca 2357  
 Db 7057 gtaccagagccttctgagcacaagaaactgacgtccacatcgttacttaggaca 7113  
 Qy 2358 gaactatatacagaacgaagtgaggatgaagaatgaagaatgaagaatgaagaatgaaga 2417  
 Db 7114 gaactatatacagaacgaagtgaggatgaagaatgaagaatgaagaatgaagaatgaaga 7173  
 Qy 2418 agaaccagagaggttaataaagccttaagcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2477  
 Db 7174 agaaccagagaggttaataaagccttaagcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 7232  
 Qy 2478 ccaagtaataagtagaagcttccatctgttccagagcgttccatctgttccagagcgttcc 2537  
 Db 7233 -----atagtaaaagatcatacttctgtgtttagggcttcttctgtctgtga 7285  
 Qy 2538 gtaagataacaggaatgagttga----- 2561  
 Db 7286 caaaacttgtggaaggggaaaaaacagccctgagatgtgctctatgttgaact 7345  
 Qy 2562 -----ctaagcttactgtgagttcgttaaaactgagc----- 2595  
 Db 7346 ctgaaactgtccttaactgttctgttctgtcgttaaccgttgcataagataaaaa 7405

Qy 2596 -----gaccatagaagaattgtatlaacatltgacagccctag 2633  
 Db 7406 gaggagaagatcgaattgtccttaacgagcccgtaagatcggtgttaccacaaatgttga 7465  
 Qy 2634 tgacctatcctaactgcaatctgttactctgtcccaagagcccaagcagatgcgactctc 2693  
 Db 7466 aacatacttctgtgacaacatgtctcccccacccgaacacatgcgcaaatgtgttaac 7525  
 Qy 2694 gggacttctttaaataatgtgtgtcccaagagcgcggcctctcgtatatttaaataatg 2753  
 Db 7526 taaacaacatttaaatgaattgtgtcccaagagcgcggcctctcgtatatttaaataatg 7585  
 Qy 2754 gtccatgagcgcggcgtctcgtatatttaaataatgtgtgtgtgtgtgtgtgtgtgtgt 2813  
 Db 7586 gt-----tctgtatatttgaataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 7630  
 Qy 2814 ttgtgaaccccaataaagcgtgtcccgatctcgcagactcgggcccagctccttaccctg 2873  
 Db 7631 ttgtaaaccccaataaagcgtgtcccgatctcgcagactcgcagcagcagcagcagcagc 7690  
 Qy 2874 cgt 2933  
 Db 7691 cgt 7750  
 Qy 2934 caaa 2938  
 Db 7751 caaga 7755

RESULT 11  
 AAT74884  
 ID AAT74884 standard; cDNA; 7892 BP.  
 AC AAT74884;  
 DT 09-FEB-1998 (first entry)  
 DE Minkature swine retrovirus cDNA.  
 DE XX  
 KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
 KW xenotransplantation; Infectious; provirus; organ transplant; donor;  
 KW activated virus; PCR; ss.  
 OS Porcine retrovirus.  
 OS XX  
 FH Key Location/Qualifiers  
 FT CDS 585..2159  
 FT /tag= a  
 FT 585..2156  
 FT /tag= b  
 FT /note= "putative GAG protein"  
 FT CDS 2307..5744  
 FT /tag= c  
 FT 2307..5741  
 FT /tag= d  
 FT /note= "putative POL protein"  
 FT CDS 5620..7536  
 FT /tag= e  
 FT 5620..7533  
 FT /tag= f  
 FT /note= "putative ENV protein"  
 PN W09721836-A1.  
 PD 19-JUN-1997.  
 PD XX  
 PF 13-DEC-1996; 96WO-0519680.  
 PF XX  
 PR 14-DEC-1995; 95US-0572645.  
 PR XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PA XX

PI	Fishman JA;
XX	
DR	WPI; 1997-332804/30
DR	P-PSDB; AAW32096-W3

PT New nucleic acid from porcine retro:viruses - used for detecting  
PT viruses in transplant or other tissue and for assessing risk of  
PT transmitting infection to graft recipient

PS Claim 22; Fig 3; 128pp; English.

CC This cDNA sequence represents a porcine retrovirus from miniature swine  
CC containing the putative coding regions for viral GAG, POL and ENV  
CC proteins. This sequence and PCR fragments generated from it  
CC (see AAT74812-774882) can be used to screen organs for the presence of  
CC porcine retroviruses prior to xenotransplantation. Transplantation can  
CC increase the likelihood of retroviral activation if intact and  
CC infectious proviruses are present. The porcine retroviral sequence can be  
CC used to generate probes to determine the level (e.g. copy number) of  
CC intact (i.e. potentially replicating) porcine provirus sequences in a  
CC strain of xenograft transplantation donors. It can be used to detect  
CC mutations, genetic lesions or viral recombinants and also to determine  
CC the histological localisation of activated retrovirus. Using Polymerase  
CC Chain Reaction DNA Quantitation (PQD) on blood mononuclear cells,  
CC infectivity titration and susceptibility testing can be performed.  
CC Ultimately animal donors without intact porcine retroviral sequences or a  
CC lower copy number of viral elements could be selected.

SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match	45.6%;	Score 1347.8;	DB 18;	Length 7892;
Best Local Similarity	74.9%;	Pred. No. 0;		
Matches 1634; Conservative	0;	Mismatches 522;	Indels 93;	Gaps 8

OY	1	tgcttttaagggttagggaaacaccccttgaaacaatttggtgtgagaccccttatgtaattgctc	60
Db	5257	tgccttttaagggttaggaacacccctgagacagatttgggtgcaccccttatgtaattctc	5316
OY	61	acgggggagaccccccccgctgtgtagaattgctctctgtataatagtgtgctgtcctt	120
Db	5317	acgggggagaccccccccatattgtctctgtacataagtgtgcgcgtgcctt	5376
OY	121	cccaagcctttgttctctataggtcgaagcgctcgtagtgggtgggtgagcaacagcgttgaac	180
Db	5377	cccaagcctttgttctctataggtcgaagcgacattgtagtgggtgagacaacagcgttgaagcc	5436
OY	181	agctcgggagagcctcactccagagagagagagacttgcgaagtccacatcgtcttccaaattg	240
Db	5437	aactcgggagagcctcactccagagagagagagacttgcagatccacacatcgtttccaaattg	5496
OY	241	gagattcagctctagttagacgcacacgcgtgtcagagaaacctcgaagactcgttggaaaggac	300
Db	5497	gagattcagctctagttagacgcacacgcgtgtcagagaaacctcgaagactcgttggaaaggcc	5556
OY	301	cttactctgtaacttttgaccacacacaaagcgctgtgaagtctgaaggaaactctcaccttga	360
Db	5557	cttactctgtaacttttgaccacacacaaagcgctgtgaagtctgaaggaaactctccaccttga	5616
OY	361	tccatcgcattcccaagcttlaagctggtgcgcacccctccgaactcgggtgtgaagaagcgaaga	420
Db	5617	tccatcgcattcccaagcttlaagctggtgcgcacccctccgaattcgggtgtgaagaagcgaaga	5676
OY	421	ctgagaatcccttaagcttcgctccatcgcgcctgtgtctcctactctaaacaataactccc	480
Db	5677	ctgagaatcccttaagcttcgctccatcgcgcctgtgtgtcctactctgtcaataactctct	5736
OY	481	cagggcagtagtagaaacgcctctatagaacagctcgaaccccatagacacttatcccttacc	540
Db	5737	cagactaaatgtatgcgtacatagagaacagcctcgaatcccatcaaaaccttatctctacc	5796
OY	541	tggctgattattagaccctctatagcgtgtgtcactgttaaatagcactcgaagcttgcctctc	600

Db	5797	tggtatctactgaccccgccacaggttatataatcaacaacacccaagggaagcctct	5856
Qy	601	aagagcaactggtgacctgaactgcattctctgcctccgatctgataaccgcgtgttaa	660
Db	5857	ttagaacactggtgacctgatactacgtttgcctgaagatccagctattc--ctagctg	5913
Qy	661	agacacactcccaacacagacgcgtagttatggttctatctgtctccacagcaag--ag	717
Db	5914	accctcaaccccgatatacctccactgctcaacgatttatglttgcgcaggaccaccaat	5973
Qy	718	aaagagaatactgtgggggtctctgggaatccctctgtagagatgagctgcgtacc	777
Db	5974	aatgaaanaaatgtcggaataatcccaagagattctttgttaacaaatggaaactgttaac	6033
Qy	778	ttcaacagatggagactgaaatgagccgactctctcccaagaccgggttaaatctccctt	837
Db	6034	tctaatgtatgataatgtggaatgtgccaactctccagcagagataaggtaagtcttctat	6093
Qy	838	gtcaaa-----ttccggcccgggcgaagtacaataatgataactataaagat	885
Db	6094	gtcaacactataccagctctgtgacaatttaattactctgacctgttagaactggaagc	6155
Qy	886	aagagcgtcccccatacagcttagattatcttaagataagtttaactgaaaaaagaaaa	945
Db	6154	ccccagtgctctcccttcaagaccatagatlaactaanaataagtttcaactgtagaagaaa	6213
Qy	946	caggaataattccaaggtgatctaaatgtgatagctctgggagatagttttta--ataa	1002
Db	6214	caagaaataatcccaaaatgtggtaaatgtgatgtctctggggaatgtatatatagtgggc	6273
Qy	1003	tatggcggggagcaggggtccactttaaccattgcctctagatatagagcgggacagaa	1063
Db	6274	tcgggtataacaacacagcctccattcttaactatctgcctctcaaaataaacaccagct--	6330
Qy	1063	ccccctgtgccaatgggagcccgataagactgtgcctgcaacaggggccccggcccttgag	1122
Db	6331	ccctcaatgtgtatagagccaataatagcgtcttgacgggtccaagaccccaaccaaaga	6390
Qy	1123	ccaccgcaataactgtccgggtgccccaattaaacctgcgtcgggcctgtgacataacagccg	1182
Db	6391	ccagacacatctct-----	6405
Qy	1183	ccctgacacagtcacactgattgtattcttaccacaacgcctagaaaactccacagtggt	1244
Db	6406	-----aacataactcttgatccaagaccacacagatccaacagcagac-----	6449
Qy	1243	ccgtttaagacagacagagactcttcctacgtctcatccagaggagctttccaagctcaac	1303
Db	6450	-----taaatgggggcaaaaacttttagcctcatccagggagcttttcaagctctctac	6504
Qy	1303	ttcacagcaacctgtgcacctctctcttctgttgcttgcctatcctccagggctccctat	1363
Db	6505	ttccacgactccagagggtctactctctctctgttgctatgtctatgactctcgggcccaacttac	6566
Qy	1363	tatagaggatggctaaagaaagaaaataatcaatgtgtacacaagagcatbagaataatgt	1422
Db	6565	tatgaaggaatggtctagaagaaggaaaattcaatgtgtacaaaagaacatbagaaccaatgc	6622
Qy	1423	acatgtgggtgtcccgaaaiaagcttaccctcaactgaagattccgggaaaggagatgcata	1483
Db	6625	acatgtgggtgtcccaaaataagacttacccttactgaggtttctctgaaaagcaccgtgcata	6684
Qy	1483	ggaaaagctcccccatactcccaacacaccttgcctataagtaactgtgtttatgagaagcc	1544
Db	6665	ggaaaaggtcccccatactcccaacacaccttgcctataagtaactgtgtttatgagaagcc	6744
Qy	1543	ttcaaaaatacgtattatagtaacctgtgttaataaagggtgggtgggcagatcaactgtgta	1603
Db	6745	tcitgaggtcaatatatctgttactctgtttatagacaggtgggtgggcagtataactgtgata	6804
Qy	1603	accctcgtgtttccacactcaagctctcaacaaatccaagaagtgttggttcaatgtgtcaa	1663
Db	6805	accctcgtgtttccacactcgtgttttaacaaaaccaaagaatttttggaatttgtgtccaa	6866

FT	/note= "putative ENV protein"
FT	3112..4686
FT	/*tag= b
FT	/note= "putative GAG protein"
FT	4871..8060
CDS	/*tag= c
FT	/note= "putative POL protein (partial)"
PN	MO9721836-A1.
XX	
PD	19-JUN-1997.
XX	
PF	13-DEC-1996; 96WO-US19680.
XX	
PR	14-DEC-1995; 95US-0572645.
XX	
PA	(GEHO ) GEN HOSPITAL CORP.
XX	
PI	Fishman JA;
DR	
XX	
XX	WPI, 1997-332804/30.
PT	New nucleic acid from porcine retroviruses - used for detecting
PT	viruses in transplant or other tissue and for assessing risk of
PT	transmitting infection to graft recipient
XX	
PS	Claim 1; Fig 1; 128pp; English.
CC	This sequence represents the purified porcine retroviral CDNA
CC	sequence of Tsukuba-1 and contains the putative coding regions for viral
CC	proteins GAG, POL and ENV. This sequence and PCR fragments generated
CC	from the sequence (see AAT74812-TY4882) could be used to screen organs
CC	for porcine retroviruses prior to xenotransplantation. Transplantation
CC	can increase the likelihood of retroviral activation if intact and
CC	infectious proviruses are present. The porcine retroviral sequence can be
CC	used to generate probes to determine the level (e.g. copy number) of
CC	intact (i.e. potentially replicating) porcine provirus sequences in a
CC	strain of xenograft transplant donor. It can be used to detect
CC	mutations, genetic lesions or viral recombinants and to determine the
CC	histological localization of activated retrovirus. Using Polymerase Chain
CC	Reaction DNA quantitation (PQD) on blood mononuclear cells, infectivity
CC	titeration and susceptibility testing can be performed. Ultimately animal
CC	donors without intact porcine retroviral sequences or with a lower copy
CC	number of viral elements could be selected.
XX	
SQ	Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;
<hr/>	
Query Match	41.7%; Score 1231.8; DB 18; Length 8060;
Best Local Similarity	69.8%; Pred. No. 0;
Matches 1873; Conservative	0; Mismatches 697; Indels 113; Gaps 11
OY	279 ctcgagacctcgttgaaaggacattatctcgtaactttgaccacacccaagcgttga
b	
1	ctcgagactcgttgaaaggcccttatctcgtaactttgaccacacccaagcgttga
OY	339 gtgcgaagaattccaccttgatcatgatcccacgttaaagcttcggcgccactccgc
OY	61 gtgcgaagaattccaccttgatcatgatcccacgttaaagcggcgccactccgc
D	
OY	399 tcgggttgagagcccgcaaaagacttagaatccccctlaagcttcgcgcctcatcg
D	
121	tcgggttgagagcccgcaaaagacttagaatccccctlaagcttcgcgcctcatcg
OY	459 ccttaacttaacaataactccccagagccagttagtaaacgccttatagacagctga
OY	181 ccttaacttgcaataaactctcaacaaatvtgtagtcgatatagagacagccta
D	
OY	519 ccataagacttaaccctactgcgtatatttacccctatagagggttcacggttaa
D	
241	ccataaaccctatctccaccttgtttaatacgcgccgcaaggtatataatalca
OY	579 tagaactcagagtglttgctcttagaggaacctgtgtgccttgaactgaatttcg









Query Match	34.2%;	Score 1009.8;	DB 22;	Length 1980;
Best Local Similarity	72.3%;	Pred. No. 1.9e-287;		
Matches 1461;	Conservative	0;	Mismatches 447;	Indels 113;
				Gaps 7;

QY	364	atgcatacccaagcttaagctcgtgcgcacactcccgactcgggtgtggagagccgaaaagctg	423
Db	1		60
QY	424	agatcccccctaagcttcgcctccatcgcgtgtgttccttaactctaaacaaataactcccg	483
Db	61		120
QY	484	gccagtagtaaacgccttataacagctcgaacccccatagaacttatacccttaactctg	543
Db	121	gttaatgtaaaagccttcttgacacagcccgaaactcccatataaacccttatctctcaactg	180
QY	544	ctgattatgacacctgtatccgggtgtctcaatgttaataagcaactccgaggtgtctccctaga	603
Db	181	ttacttaactggaactccgatacaggtataataataacagcaactcaaggaggtgagctctctg	240
QY	604	ggcactctgtgtgcctgaacatgcattctctgcctccgatgttaacccccgttta-----	658
Db	241	gggactctgtgtgcctgaatataatgtctgtccttgtaacgtataccctggtctcaatgac	300
QY	659	-aaagcacactcccaacctagatccgtatgtatgtgtctatctgtcccgagg---caca	714
Db	301	caggcacacaccccccgaatbtaactcgtgtcgttccggttttaacgttttgcgcagggcccca	360
QY	715	ggagaaaggaaataactgtgtgggtgtggtctgtgggaatcctctctgtagaagatgtgagctgcgtc	774
Db	361	aataatggaagaatattgtgtgaaatccctcagaagctctcttcttgcaagaacaatgtgagctggtla	420
QY	775	acctccaacgatlbgagactggaatagtcgcgaactctctctccagagccgggtataactctcc	834
Db	421	actcttaatgatlbggaattgtgaatctgcagctctctccagacagagtgtaagttaacct	480
QY	835	tttgt-----caatlc	845
Db	481	tttgttaacaactcctacagttataatcaatttaattatgtgcaltggaatgtgaaagat	540
QY	846	cggcccggtgcaagtcaaatatgatataaactataagaatagagctgtcccccataga	905
Db	541	tggcaacagcggtgtacaaaaagatg-tacgaaataagaacaataagctgtcatctgttga	599
QY	906	cttagattactcaagaataagttctcaactgtgaaaaaggaaacagggaaaaatattcaaaagt	965
Db	600	cttagattactcaaaaaataaagttctcaactgtgaaaaaggaaacaaataatttcaaaagt	659
QY	966	gataatgtgataagcttggggaatggttttttaatakatatggtcggtggagagc---agggtc	1022
Db	660	gttaaatgtgataagcttvggggaatagttgtaactatagaggtcttvggagaaagaaagatc	719
QY	1023	caatttaaccattctgcctttagatagatagaacgggtgacaaaccccttggcaatgtggacc	1082
Db	720	tgttctgactattgcgcctcagataagaaatcagatgtgaactccgtgtgtatagaacc	779
QY	1083	cgataaagtactgtgtgtaacaggggtcccccggccctgtgagccacgcgcataacttgcggt	1142
Db	780	aaataaagggttcttggcggacaagaagcctccatccaaagaa-----	819
QY	1143	gcccacaatlaacctgcgcgcgcgtgacataacagaacggcgttagaagcttaccatgtg	1202
Db	820	-----cagagggcatctcccaaccctctatatacaataaactcttgg	863
QY	1203	attgattcttaacaaacacgcgtcagaactcccccaggtgttccgttlaagacagacagag	1262
Db	864	atcagttcccaactagcctaacaatca-----ctattaaacaggggcgaa	908
QY	1263	actcttcagcttcatccagggagcttccaaacgacataacacacacacggacgacttgtgcac	1322
Db	909	acttttaacctatccagggagctttttcaagctcttaactccacgactccagactccaggtctac	968
QY	1323	ttctctctgttggcttctgtatccatccaaaggctcctcttattatgagggatgtgttaaaga	1382

Db	969	ctctctcgttggtccttgcttagcttcgtggcccaaccttactatgagggagctgtagag	10282
QY	1383	aagaataatcgaatgtgacccaagaagcatalgaataatcgaatctgacatctggggctcccgaaata	14427
Db	1029	agggaataatcgaatgtgacaaaaggagacatagagccaatgtactctggggtcccaataaa	10881
QY	1443	gcttaacctcactgaagttctccgggaaggagacatgcataaganaaaagctcccccata	15023
Db	1089	gcttaacctcactgaagttctcggaaaaggacactgacatagatggtctcccccata	11482
QY	1503	ccaacaaccttgctatagtagctgtggttatatgagcaggtcccaagaataatcagatattg	15676
Db	1149	ccaacaaccttgtaacccaacacatgtagcctttaaactgaacctctgagagtcagatctg	12080
QY	1563	acctgttataacaggtgtgtggcagatgcatactgggtttaaccccttggttccaact	16222
Db	1209	acctgttataacaggtgtgtggcagatgcatactgggtttaaccccttggttccaact	12686
QY	1623	agcttccaaccaatccaagaattgtgtgtcagatgtgccaaatgttccccaggttacta	16821
Db	1269	agtttccaaccaataaagaactttggtgtatgtgtccaaatgttccccaggttacta	13282
QY	1683	ccatccctgaaggaagtgtgtctctgtatgaatgatcctcgttataacccaacgaacaa	17422
Db	1329	ctatcccgaaaagaacatctctgtatgatatatgactatagatataatcgcacaaaga	13888
QY	1743	accggttatcccttacccttaactgtgaatagctcggatattaggaagcgcgttgggtgtagaac	18022
Db	1389	gcccatatcccttgacacttagctgtgaatagctcggatattggaagtgtgtgcaggtgtggagac	14488
QY	1803	agggaacagctgtccctgtatccacagaaccacacgcgcgtcagagaaggaactgtgtgagctaca	18622
Db	1449	agggaacagctgtcccttaataccaggaacccgcgaacacgcgcgtcagagaaggaactgtgaaacctaca	15088
QY	1863	tggtgcacatgacagaagatctccgagccttaaaaggagctctgtttagcaacctagaagaag	19222
Db	1509	tcgaattgtaacggagaataatctccaagccttagaanaaatctgcagtaaccttgtaggaatc	15688
QY	1923	cctgacttcttctgtctgaagtgtgtcttaagaacccggagaggatagatctgcgttctc	19822
Db	1569	cctaacctccttactctgaagtgtgtcttaccagaacagaagggtgttagactcgttattctc	16282
QY	1983	aagaagaagtggtgtatgtgtgcagccttaaaagaagaatgtgtcctctatgttagatcactc	20422
Db	1629	aaaagaagaaggaattatgtgttagcctttaaaggaggaatgtgttcttctatgttagatcactc	16888
QY	2043	aggagccatccagagatctcgaatgaacaagctctgaanaaaagttagagaagcgtctgaagga	21022
Db	1689	aggggccatccagagatctcgaatgaacaagctctgaagaaggttggggaagcgtctgaagga	17488
QY	2103	aagaagaagctgacccaagggtgtgttgaagaagtgttcaaacagctcctcttgatgacac	21622
Db	1749	aaaagaagaactacttaaaagggtgtgttgaagaagtgttcaaacaggtcctccttggtgtac	18088
QY	2163	ccctgtcttcgcctctgacggggcccttagtagtctcgtccctgttactataagtgtggcc	22222
Db	1809	ccctgtcttcgcctctgacgaagacccttaataagctccctccctgttactccaagtgtggcc	18688
QY	2223	ttgcttaataatagagtttctgtccttctttagagaacagatgtgagctgagctcagatcat	22822
Db	1869	atgtattattataacaagaagtattgtccttcaatttagagaacgaataaagtgtgagctcagatcat	19282
QY	2283	ggtacttaggcacaacagtlaccaagcctgcgtctagcagagaag	23232
Db	1929	ggtacttagacaacagtlaccaagaagccgcgtctagcagagaag	19692
RESULT 15			
AAZ51087			
ID AAZ51087 standard; DNA; 2000 BP.			
XX AAZ51087:			
GC			



Db	1363	acactgcagtaagctcgtcagctcgtgagatcaacagctgtgagaaacaggaactgcagct	1422
QY	1816	ctgtacacaggaaccaacagcagctcagagaagaagacttgtgtgagctacatcgccgcatgaca	1875
Db	1423	ttagttaacaggaacctcagcagctctagaacaagagacttagtaacctacatcaaatgtgaca	1482
QY	1876	gaagatctccgagcgccttaaaagagctctgttgaacaacctgagaagagctccgaactcctgt	1935
Db	1483	ggaatctccagagccttaaaaaaactctgtccagtaacctggaataaacctcaacctctta	1542
QY	1936	lctgaagctgtgtctacagaaacccgagagggattagatctctgttcttaagaaaggttgag	1995
Db	1543	ctctgaagtagttctacagaaataaaagggttagattattattcttaaaaaaagaaga	1602
QY	1996	ttatgtgcagccttaaaagaagaagtctgtcctctatgtatagatcaactcagagagcatcaga	2055
Db	1603	ttatgtgttaaccttaaaaggagaaatgctgttcttctatgtatcaatcttcagagggcatcaga	1662
QY	2056	gactccaatgacaagctctaaaaaaaaggtttagaagagcgctgcgaaggagaaaggagcgctgac	2115
Db	1663	gactccactgacaacagctttaaaaaaaggttggagaaacgttcgaagggaaaaggaaaacttac	1722
QY	2116	-caggggtgtgtcttgaagagatgcttcaacacagctctccttgaatgacacccgctctctgc	2174
Db	1723	tcaagatgtgtttaaaagatggttcaacacagctctcctctgtgtgtgtgtgtgtgtgtgtgtgt	1782
QY	2175	ctctgaagcgggccctcagtagtctgtcctgtcctgttacttacaagctgttgccctgtctaataa	2234
Db	1783	tttaacagagaccccttaataatgctctcctcctctgttacttacaagaatgttggtccatgtatataaa	1842
QY	2235	tagattgtgtgcttgttgaagaagatgtgagtcagatccagatcacaatgcttacttaagca	2294
Db	1843	caagttatattgctctcaatctagaaaaaagaataatgtcagatccagatcacaatgcttacttagaca	1902
QY	2295	acagttacaaagcctctctgagccaagagaaaactgaaccttagtcccttccacgttcttaaga	2354
Db	1903	acagttacaaacccatctca--acagaggaacgtgagccgtacgtcttaccacgttcttaaga	1959
QY	2355	tttaaaactattaaacagaacagagatgggggaatgtgaagaat	2395
Db	1960	ttaaactattaaacagaagagaagatgggggaatgtgaagaat	2000

```
Search completed: February 24, 2002, 03:05:27
Job time: 19493 sec
```



OM of: US-09-171-553b-10 to: N\_Geneseq\_1101:\* out\_format : pfs  
 Date: Feb 24, 2002 10:24 AM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet p2n.model -DEV=xlh  
 -O=/cpn2.1/USPTO\_spool/US09171553/runat\_22022002\_122826\_11203/app\_query.fasta\_1.3312  
 -DB=N\_Geneseq\_1101 -QFMT=fastcap -SUFFIX=feb22std.rng  
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000  
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45 -DOCALLIGN=200  
 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
 -OUTFMT=pfs -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09171553 -CGN1\_1.526 -NCPU=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-171-553b-10  
 Query length: 657  
 Database: N\_Geneseq\_1101:\*  
 Database sequences: 930621  
 Database length: 42862619  
 Search time (sec): 338.200000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	! Documentation
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09703		3533.00	4463.44	1974	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV82749		3484.00	4395.62	1.5e-240	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699		3445.50	4338.21	9.2e-237	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09700		3444.50	4336.93	1.5e-233	
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAV82748		2631.50	3318.07	1.7e-233	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67022		2546.50	3207.15	9.7e-177	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67018		2546.50	3212.47	7.4e-171	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77726		2537.50	3187.93	1.7e-169	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV77727		2537.50	3187.85	1.7e-169	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV67023		2514.50	3158.05	7.9e-168	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV74811		2402.00	3015.16	7.2e-160	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77725		2402.00	3015.16	7.2e-160	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV74884		2395.00	3006.50	2.2e-159	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV77727		2395.00	3006.20	2.3e-159	
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAV51086		2169.00	2733.68	3.4e-144	
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAV46343		1558.50	1958.90	4.9e-101	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV90691		1544.50	1930.25	1.9e-99	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV80076		1526.00	1918.04	9.3e-99	
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAV46344		1522.50	1913.25	1.7e-98	
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAV45540		1522.50	1899.87	9.5e-98	
/SID22/gcgdata/geneseq/geneseq/NA1994.DAT:AAV073731		1506.50	1893.37	1.0e-97	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV66507		1506.50	1893.57	2.1e-97	
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAV89080		1488.00	1870.14	4.3e-96	
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAV69750		1478.00	1843.40	1.3e-94	
/SID22/gcgdata/geneseq/geneseq/NA1995.DAT:AAV81042		1329.00	1654.21	4.6e-84	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV80058		1329.00	1654.21	4.6e-84	
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAV07063		1329.00	1654.21	4.6e-84	
/SID22/gcgdata/geneseq/geneseq/NA1994.DAT:AAV508700		1315.00	1636.46	4.5e-83	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30412		1288.50	1616.06	6.1e-82	
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAV30413		1285.50	1613.36	8.7e-82	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30414		1285.50	1612.06	1.0e-81	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30415		1285.50	1611.72	1.1e-81	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV07063		1285.50	1610.55	1.2e-81	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30406		1284.00	1609.95	1.3e-81	
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV30407		1284.00	1609.76	1.4e-81	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV90690		1284.00	1600.33	4.6e-81	
/SID22/gcgdata/geneseq/geneseq/NA1995.DAT:AAV08452		1284.00	1597.14	6.9e-81	
/SID22/gcgdata/geneseq/geneseq/NA1996.DAT:AAV07374		1284.00	1597.14	6.9e-81	
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV05061		1284.00	1597.14	6.9e-81	

/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV42715 + 1284.00 1597.14 6.9e-81  
 /SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV19551 + 1284.00 1597.14 6.9e-81  
 /SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAV15346 + 1284.00 1597.14 6.9e-81  
 /SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAV28143 + 1284.00 1597.14 6.9e-81  
 /SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAA40297 + 1284.00 1597.14 6.9e-81

seq\_name: /SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09703

seq\_documentation\_block:

ID AAV09703 standard; DNA; 1974 BP.  
 AC AAV09703;  
 DT 20-MAY-1998 (first entry)  
 XX Porcine retrovirus Raji clone ENV DNA.  
 DE Porcine retrovirus; PoEV; ENV protein; envelope protein; vaccine;  
 KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
 XX Porcine retrovirus.  
 OS  
 FH Key Location/Qualifiers  
 CDS 1..1974  
 FT /\*tag= a  
 FT /product= ENV protein  
 FT /note= "envelope protein"  
 XX  
 PN W09740167-AL.  
 XX  
 PD 30-OCT-1997.  
 XX  
 PF 18-APR-1997; 97WO-GB01087.  
 XX  
 PR 10-FEB-1997; 97GB-0002668.  
 PR 19-APR-1996; 96GB-0008164.  
 XX  
 PA (IMUT-) IMUTRAN LTD.  
 PA (QONE-) Q-ONE BIOTECH LTD.  
 XX  
 PI Galbraith DN, Haworth C, Lees GM, Smith KT;  
 WP1: 1997-535851/49.  
 DR P-PSDB; AAV39274.  
 XX  
 PT Polynucleotide encoding porcine retrovirus expression product -  
 PT useful to develop products for use in vaccines, diagnosis and  
 PT xeno-transplantation  
 PS Claim 6; Fig 4; 69pp; English.  
 XX  
 CC This sequence encodes the porcine retrovirus envelope (ENV) protein  
 CC isolated from the human cell line Raji. Such viral proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.  
 XX  
 SQ Sequence 1974 BP; 534 A; 489 C; 483 G; 468 T; 0 other;

alignment\_scores:  
 Quality: 3533.00 Length: 657  
 Ratio: 5.377 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-171-553B-10 x AAV09703 ..  
 Align seg 1/1 to: AAV09703 from: 1 to: 1974  
 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyCylupr 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 1 ATGCATCCCACTTAAGCTGGCGCCACCTCCCGACTCGGGGTGGAGGCC 50

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 |||||  
 51 GAAAGACTGAGAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCCTTA 100  
 |||||  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSer 50  
 |||||  
 101 CTCTAACATAACTCCCGAGCCAGTAGTAACGGCTTATAGACAGCTCG 150  
 |||||  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||||  
 151 AACCCCATAGACCTTTATCCCTACCTGGCTGGTATATTGACCCGTGATAC 200  
 |||||  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 |||||  
 201 GGGTGTCACTGTAAATAGACACTCGAGGTGTGCTCCTAGAGGCACCTGGT 250  
 |||||  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 |||||  
 251 GGCCCTGAACCTGATTTCTGCCTCCGATTGATTAAACCCGCTGTAAAGC 300  
 |||||  
 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 |||||  
 301 ACACCTCCCACTAGTCCGTAGTTATGGTTCTATTGCTGCCAGGCAC 350  
 |||||  
 117 rGluLysGluLysTyrCysGlySerGlyGluSerPheCysArgArgT 134  
 |||||  
 351 AGAAGAAAGAAATACTGTGGGGTCTGCGGGAATCTCTTGTAGGAGAT 400  
 |||||  
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 |||||  
 401 GGAGCTGCTGCTCACTCCAAACGATGGAGACTGGAAATGGCCCATCTCTC 450  
 |||||  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
 |||||  
 451 CAGGACCGGGTAAATTTCTCTTTGTCATATCCGGCCGGGCAAGTACAA 500  
 |||||  
 167 smetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
 |||||  
 501 AATGATGAACATATATAAGATAGAGCTGCTCCCATCAGACTAGATT 550  
 |||||  
 184 yrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 200  
 |||||  
 551 ATCTAAAGATAAGTTTCACTGAAAGAGGAAACAGGAAATATTCAAAG 600  
 |||||  
 201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyG 217  
 |||||  
 601 TGGATAAATGGTATGAGCTGGGGAATAGTTTTTATAAATATGCGGGGG 650  
 |||||  
 217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 234  
 |||||  
 651 AGCAGGTGCCACTTTAACCATTCGCTTAGGATAGACGGGGGACAGAAC 700  
 |||||  
 234 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250  
 |||||  
 701 CCCCTGTGGCAATGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCG 750  
 |||||  
 251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267  
 |||||  
 751 GCCCTGGAGCCACCGCAATCTGCCGGTGCCCAATTAACCTCGCTGCG 800  
 |||||  
 267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProT 284  
 |||||  
 801 GCCTGACATACACAGCGGCTAGCAACAGTACCACCTGGATTGATTCCTA 850  
 |||||  
 284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300  
 |||||  
 851 CCAACAGCCCTAGAACTCCCGAGGTGTTCTGTTAGACAGGACAGAGA 900  
 |||||  
 301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317  
 |||||  
 901 CTCTTCAGTCTCATCCAGGGAGCTTTCCAAAGCCATCAACTCCACGACCC 950  
 |||||

317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyrT 334  
 |||||  
 951 TGATGCCACTTCTTTGTTGGCTTTGTCTATCCTCAGGGCCTCCTTATT 1000  
 |||||  
 334 yrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350  
 |||||  
 1001 ATGAGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGCATAGA 1050  
 |||||  
 351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSe 367  
 |||||  
 1051 AATCAATGTACATGGGGTCCCGAATAAGCTTACCCTCAGTCAAGTTTC 1100  
 |||||  
 367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 384  
 |||||  
 1101 CGGAGAGGGACATGTCATAGAAAGCTCCCATCCCAACACACCTTT 1150  
 |||||  
 384 ystYrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400  
 |||||  
 1151 GCTATAGTACTGTGTTTATGAGCAGGCTTCAGAAATCAGTATTAGTA 1200  
 |||||  
 401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417  
 |||||  
 1201 CCTGGTTATAACAGGTGGTGGCATGCAATACTGGGTAAACCCCTGTGT 1250  
 |||||  
 417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434  
 |||||  
 1251 TTCACCTCAGTCTTCAACCAATCCAAAGATTGTTGTCTATGGTCCAAA 1300  
 |||||  
 434 leValProArgValTyrTyrHisProGluGluValValLeuAspGluTyr 450  
 |||||  
 1301 TCGTCCCCGAGTGTACTACCATCTCGAGAGGTGCTTGTATGAATAT 1350  
 |||||  
 451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467  
 |||||  
 1351 GACTATCGGTATAACCGACCAAAACAGAACCCGTATCCTTACCCTAGC 1400  
 |||||  
 467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrAlaA 484  
 |||||  
 1401 TGTAAATGCTCGGATTAGGACGGCCGTTGGCGTAGGAACAGGACAGCTG 1450  
 |||||  
 484 laLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHis 500  
 |||||  
 1451 CCCTGATCAGGACCCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACAT 1500  
 |||||  
 501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517  
 |||||  
 1501 GCGGCCATGACAGAGATCTCCGAGCCTTAAAGAGTCTGTTAGCAACCT 1550  
 |||||  
 517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgG 534  
 |||||  
 1551 AGAAGAGTCCCTGACTTCTTGTCTGAAGTGGTTCTTACAGAACCCGAGGG 1600  
 |||||  
 534 lyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLys 550  
 |||||  
 1601 GATTAGATCTGCTGTTCTTAAGAGAGGTGGGTATGTGACGCCCTAAAA 1650  
 |||||  
 551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567  
 |||||  
 1651 GAAGAATGTTGCTTCTATGTAGTCACTCAGGAGCCATCAGAGACTCCAT 1700  
 |||||  
 567 tAsnLysLeuArgLysLysLeuArgArgArgGluArgGluAlaA 584  
 |||||  
 1701 GAACAGCTTAGAAAAAGATTAGAGAGCGTCGAAAGGGAAGAGAGGCTG 1750  
 |||||  
 584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600  
 |||||  
 1751 ACCAGGGTGGTTGAAGGATGGTTCAACAGGCTCCTCTGGATGACCAACC 1800  
 |||||  
 601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuTh 617  
 |||||  
 1801 CTGCTTCTGCTCTGACGGGGCCCTAGTAGTCTGCTGCTCTTACTTAC 1850  
 |||||  
 617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgv 634  
 |||||



```

|||||
1851 AGTTGGGCTTGTCTAAATTAATAGTGTGTGCTTTGTTAGAGAACGAG 1900
634 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
|||||
1901 TGAGTGCAGTCAGATCATGTACTTAGGCAACAGTAGCAAGGCGCTTCTG 1950
651 SerGlnGlyGluThrAspLeu 657
|||||
1951 AGCCAAGGAGAACTGACCTC 1971

```

seq\_name: /SDS2/gcddata/geneseq/geneseqn/NA1999.DAT.AAV82749

seq\_documentation\_block:

ID AAV82749 standard; DNA; 3482 BP.

AC AAV82749;

XX 25-FEB-1999 (first entry)

DT Pig endogenous retrovirus (PERV)-B envelope (env) gene region.

DE Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
KW probe; primer; detection; retrovirus; human tissue; xenotransplant;  
KW primary porcine tissue; human cell line; porcine cell line; ss.  
XX

OS Pig endogenous retrovirus.

XX Key Location/Qualifiers

FT CDS 911...2884

FT /\*tag= a

FT /product= envelope\_protein

XX WO9853104-A2.

XX 26-NOV-1998.

XX 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Stoye JP, Weiss RA;

XX WPI; 1999-045324/04.

XX P-PSDB; AAW85452.

XX Newly isolated nucleic acid probe capable of hybridising to either  
PT the PERV-A or PERV-B env gene - useful in the detection of  
PT retroviruses, and their subtypes, in a sample of porcine/human  
PT tissue

XX Claim 3; Page 21-23; 36pp; English.

XX The present sequence encodes a pig endogenous retrovirus (PERV)-B  
CC envelope protein. PERV exists in two different subtypes, PERV-A and  
CC PERV-B. The differences are reflected in sequence divergence in the  
CC envelope genes. Probes and primers can be derived from the envelope  
CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
CC a method to detect retroviruses in a sample of porcine/human tissue,  
CC particularly primary porcine tissue and human cell lines that have been  
CC cultivated in the presence of a porcine cell line, or human tissue from  
CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
CC one of the PERV env genes can also be determined.

XX Sequence 3482 BP; 927 A; 854 C; 867 G; 834 T; 0 other;

alignment\_scores:

Quality: 3484.00

Ratio: 5.319 Length: 657

Percent Similarity: 99.696 Gaps: 0

Percent Identity: 98.478

alignment\_block:

US-09-171-553B-10 x AAV82749 ..

Align seg 1/1 to: AAV82749 from: 1 to: 3482

```

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
|||||
911 ATGCATCCACAGTTAAGTGGCCGACCTCCGACCTCGGGTGGAGAGCC 960
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||
961 GAAAGAGACTGAGAAATCCCTTAAGCTTGGCTCCATCGCTGCTCTCTTA 1010
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuLeuLeuSerSer 50
|||||
1011 CTCTAACAAATACTCCCGAGGCGAGTAGTAAGCGCTTATAGACAGCTCG 1060
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||
1061 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGTATAC 1110
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
1111 GGGTGTCACTGTAAATAGCAGCTCGAGGTGTGCTCTAGAGGCACCTGGT 1160
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
1161 GGCTTGAATGCAATTTCTGCCGCTCCGATTGATTAAACCCGCTGTTAAAGC 1210
101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
1211 ACACCTCCCAACCTAGTCGATGATGATGGTCTATTGCTGCCAGGCAC 1260
117 rGluLysGluLysTyrCysGlySerGlyGluSerPheCysArgArgT 134
|||||
1261 AGAGAAAGAGAAATACTGTGGGGTCTCGGGAAATCCTTCTTAGGAGAT 1310
134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysSerLeu 150
|||||
1311 GGAGCTGCGTCACTCCCAACGATGGAGACTGGAAATGGCCGATCTCTCTC 1360
151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIy 167
|||||
1361 CAGGACCGGGTAAATTTCTTCTGTCATTCGGCCCGGGCAAGTACAA 1410
167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184
|||||
1411 AGTGATGAAACTATATAAAGATAAGAGCTGCTCCCATCAGACTTAGATT 1460
184 yrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLys 200
|||||
1461 ATCTAAAGATAAGTTTCACTGAAAAGGAAACAGGAAATATTCAAAG 1510
201 TrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGly 217
|||||
1511 TGGATAAATGGTATGAGCTGGGGAATAGTTTTTATAATATGGCGGGG 1560
217 yAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGluP 234
|||||
1561 AGCAGGGTCCACTTTAAACATTTCGCTTAGGATAGAGACGGGACAGAAC 1610
234 roProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPro 250
|||||
1611 CCCCTGTGCGACTGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCGG 1660
251 AlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuAr 267
|||||
1661 GCCCTGGAGCCACCGCATAACTTGGCGTGCCCAATTAACCTCGCTCGG 1710
267 gProAspIleThrGlnProProSerAsnSerThrThrGlyLeuLeuProT 284
|||||
1711 GCCTGACATAAACACGCCCTTAGCAACGGTACCAGTGGATTGATTCTCTA 1760

```

```

284 hrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnArg 300
|||||
1761 CCAACACGCTAGAAACTCCCGAGGTGTTCTGTTAAGACAGGACAGAGA 1810
|||||
301 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspPr 317
|||||
1811 CTCCTTCAGTCTCATCCAGGGAGCTTCCACGCCATCAACTCCACGAGCC 1860
|||||
317 oAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 334
|||||
1861 TCATGCCACTCTCTCTGTTGCTTGTCTATCCTCAGGGCTCCTTATT 1910
|||||
334 yGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArg 350
|||||
1911 ATGAGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGACATAGA 1960
|||||
351 AsnGlnCysThrTrpGlySerArgAsnLysLeuThrGluValSe 367
|||||
1961 AATCAATGTACATGGGGTCCGAANTAAAGCTTACCCCTCACTGAAGTTTC 2010
|||||
367 rGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuC 384
|||||
2011 CGGGAAGGGACATCATAGGAAGCTCCCCATCCCAACACACCTTT 2060
|||||
384 ysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVal 400
|||||
2061 GCTATAGTACTGTGTTTATGAGCAGGCTCAGAAATCAGTATTAGTA 2110
|||||
401 ProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVa 417
|||||
2111 CTGTTTAAACAGTGTGGCATGCAATCTGGTTAAACCCCTGTGT 2160
|||||
417 lSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGlnI 434
|||||
2161 TTCCACCTCAGTCTCAACCAATCCAAAGATTCTGTGTCATGTGCCAAA 2210
|||||
434 leValProArgValTyrTrpHisProGluGluValValLeuAspGluTyr 450
|||||
2211 TCGTCCCCCGAGTGTACTACCATCTCGAGAGTGGTCTTTCATGAATAT 2260
|||||
451 AspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAl 467
|||||
2261 GACTATCGGTATACCGACCAAAAGAGAACCCGTATCCCTTACCTAGC 2310
|||||
467 aValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaA 484
|||||
2311 TCTAATGCTCGGATTAGGACGGCGTGGCGTAGGAACAGGACAGCTG 2360
|||||
484 laLeuIleThrGlyProGlnGlnLeuLysGlyLeuGlyGluLeuHis 500
|||||
2361 CCTGTATCACAGGACACAGCAGCTAGAGAAAGGACTTGTGAGCTACAT 2410
|||||
501 AlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnLe 517
|||||
2411 GCGGCCATGACAGAGATCTCGAGCCCTAGAGAGTCTGTACCACT 2460
|||||
517 uGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 534
|||||
2461 AGAAGAGTCCCTGACTCTTGTCTGAAGTGGTCTACAGAACCGGAGGG 2510
|||||
534 lyLeuAspLeuLeuPheLeuArgGluGlyLeuCysAlaAlaLeuLys 550
|||||
2511 GATTAGATCTGCTGTTTCTAAGAGAGGTGGGTATGTCAGACCTTAAAA 2560
|||||
551 GluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMe 567
|||||
2561 GAAGAATGTGCTTCTATGTAGTACTCACTCAGAGCCATCAGACTCCAT 2610
|||||
567 tAsnLysLeuArgLysLysLeuGluArgArgArgGluAlaA 584
|||||
2611 GAGCAAGCTTAGAGAAAGGTTAGAGAGCGTCGAAGGGAAGAGAGGCTG 2660

```

```

584 spGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 600
|||||
2661 ACCAGGGTGTGTTGAAGGATGTTCAACAGGTCTCTTGGATGACCACC 2710
|||||
601 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617
|||||
2711 CTGCTTCTGCTCTGACGGGACCCCTAGTAGTCTCTGCTCTGTACTTAC 2760
|||||
617 rValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgV 634
|||||
2761 AGTGGGCCCTGCTTAATTAATAGTTGTTGCTTTGTAGAGAACGAG 2810
|||||
634 alSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeu 650
|||||
2811 TGAGTCAGTCCAGATCATGTACTTAGGCACAGTACCAAGGCCTTCTG 2860
|||||
651 SerGlnGlyGluThrAspLeu 657
|||||
2861 AGCCAGGAGAAACTGACCTC 2881
|||||
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAV09699
seq_documentation_block:
ID: AAV09699 standard; DNA; 8196 BP.
XX
AC: AAV09699;
XX
DT: 19-MAY-1998 (first entry)
XX
DE: Porcine retrovirus DNA encoding, GAG, POL and ENV.
XX
KW: Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;
vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.
XX
OS: Porcine retrovirus.
XX
FH: Key Location/Qualifiers
CDS 576..2126
FT /*tag= a
FT /product= GAG protein
FT /note= "viral core polypeptide"
CDS 2143..5733
FT /*tag= b
FT /product= POL protein
FT /note= "polymerase peptide sequence as given in
specification"
CDS 5606..7576
FT /*tag= c
FT /product= ENV protein
FT /note= "envelope protein"
XX
PN: WO9740167-A1.
XX
PD: 30-OCT-1997.
XX
PF: 18-APR-1997; 97WO-GB01087.
XX
PR: 10-FEB-1997; 97GB-0002668.
PR: 19-APR-1996; 96GB-0008164.
XX
PA: (IMUT-) IMUTRAN LTD.
PA: (QONE-) Q-ONE BIOTECH LTD.
XX
PI: Galbraith DN, Haworth C, Lees GM, Smith KT;
WPI; 1997-535851/49.
XX
DR: WPI; 1997-535851/49.
XX
PT: Polynucleotide encoding porcine retrovirus expression product -
useful to develop products for use in vaccines, diagnosis and
xeno-transplantation
XX
PS: Claim 4; Fig 2; 69pp; English.
XX

```

CC This DNA sequence encodes the porcine retrovirus (PoEV) virion core  
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These  
 CC proteins can be used to develop viral vaccines, antisense nucleic acids,  
 CC ribozymes and other antiviral agents. They can also be used in  
 CC xeno-transplantation technology and as diagnostic tools.  
 XX

SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

## alignment\_scores:

Quality: 3445.50 Length: 658  
 Ratio: 5.276 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 98.936

## alignment\_block:

US-09-171-553B-10 x AAV09699 ..

Align seg 1/1 to: AAV09699 from: 1 to: 8196

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 |||||  
 5606 ATGCATCCACGTTAAGCYGGGCCACCTCCGACTCGGGTGGAGAGCC 5655  
 17 OlysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 |||||  
 5656 GAAAGACCTGAGAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTA 5705  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 |||||  
 5706 CTCTACATAACTCCCGAGCCAGTAGTAAGCGCTTATAGACAGCTCG 5755  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||||  
 5756 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCTGATAC 5805  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 |||||  
 5806 GGGTGTCACTGTAATAGACACTCGAGGTGTCTCTCTAGAGCACCTGGT 5855  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 |||||  
 5856 GGCTGAACCTGCATTTCTCCCTCCGATGATTAAACCCGCTGTTAARAGC 5905  
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 |||||  
 5906 ACACCTCCCAACCTAGTCGTAGTTATGGTTCTATTGCTGCCAGGCAC 5955  
 117 rGluLysGluLysTyrCysGlySerGlySerGlyGluSerPheCysArgT 134  
 |||||  
 5956 AGAGAAAGAGAAATACGTGGGGGTCTGGGGAATCCTTCTGTAGGAGAT 6005  
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 |||||  
 6006 GGAGCTGCCTCACTCCACGATGAGACTGGAATGGCCGATCTCTCTC 6055  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLy 167  
 |||||  
 6056 CAGGACGGGTAAATTCCTTTGTCAATTCGGCCCGGGCAAGTACAA 6105  
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184  
 |||||  
 6106 AATGATGAACATATATAAGATAAGAGCTGCTCCCATCAGACTTAGATT 6155  
 184 yrLeuLysIleSerPheThrGluLys.GlyLysGlnGluAsnIleGlnLy 200  
 |||||  
 6156 ATCTAAAGATAATTTTCACT...GAAGAGAAACAGGAAATATTCAAAA 6202  
 200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyG 217  
 |||||  
 6203 GTGGATAAATGATAGCTGGGGAATAGTTTTTATTA.TATGGCGGGG 6251  
 217 lyAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233  
 |||||

6252 GAGCAGGGTCCACTTTAAACCATTCGCCCTTAGGATTAGACGCGGGACAA 6301  
 234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250  
 |||||  
 6302 CCCCTCTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCC 6351  
 250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267  
 |||||  
 6352 GGCCCTGGAGCCACCGCATAACTTGCCTGGTCCCAATTAACTCGCTGC 6401  
 267 rgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283  
 |||||  
 6402 GGCCTGACATAACACAGCCGCTTAGCAACAGTACCACCTGGATTGATTCT 6451  
 284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300  
 |||||  
 6452 ACCAACACGCTTAGAACTCCCGAGGTGTCTCTGTTAAGACAGGACAGAG 6501  
 300 gleuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAspP 317  
 |||||  
 6502 ACTCTTCAGTCTCATCCAGGGAGCTTTCACAGCCATCAACTCCACGACC 6551  
 317 roAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333  
 |||||  
 6552 CTGATGCCACTTCTTCTTGGCTTTGTCTATCTCAGGGCTCCTTAT 6601  
 334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisAr 350  
 |||||  
 6602 TATGAGGGGATGGCTAAAGAAAGAAATTCATGTGACCAAGAGACATAG 6651  
 350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValS 367  
 |||||  
 6652 AAATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTT 6701  
 367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383  
 |||||  
 6702 CCGGAGGGGACATGATAGAAAGCTCCCCATCCCAACACACTT 6751  
 384 CysTrpSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400  
 |||||  
 6752 TGCTATAGTACTGTGTTTATGACGAGCCCTCAGAAATCAGTATTAGT 6801  
 400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417  
 |||||  
 6802 ACCTGGTTATAACAGGTGGTGGGCATCAATACTGGGTAAACCCCTGTG 6851  
 417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433  
 |||||  
 6852 TTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCAA 6901  
 434 lIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTy 450  
 |||||  
 6902 ATCGTCCCCCAGGTGTACTACCATCCCTGAGGAAGTGGTCTTGATGAAT 6951  
 450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467  
 |||||  
 6952 TGACTATCGGTATAACCGACCAACAAAGAGAACCCGTATCCCTTACCCTAG 7001  
 467 laValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483  
 |||||  
 7002 CTGTAATGCTGGATTAGGACGCGCTTGGCGTAGGAACAGGACAGCT 7051  
 484 AlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuH 500  
 |||||  
 7052 GCCCTGATCACAGGACCACAGCAGCTAGAGAAAGGACTTGTGTGAGCTACA 7101  
 500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517  
 |||||  
 7102 TGGCGCCATCACAGACAGATCTCCGAGGCTTAAAGAGAGTCTGTAGCAACC 7151  
 517 euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533  
 |||||  
 7152 TAGAAGAGTCCCTGACTTCTTTGTCTGAGTGGTCTTACAGAACCGGAGG 7201

534 GlyLeuAspLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuIly 550  
 |||||  
 7202 GGATTAGATCTGCTGTTCTTAAGAGAAGGTGGGTTTATGTCAGCCTTAA 7251  
 |||||  
 550 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567  
 |||||  
 7252 AGAAGATGTTGCTCTCTATGTAGATCACTCAGAGCCATCAGAGACTCCA 7301  
 |||||  
 567 etAsnLysLeuArgLysLysLeuGluArgArgArgGluArgGluAla 583  
 |||||  
 7302 TGAACAAGCTTAGAAAAAGTTAGAGAGCGGTCGGAAGGGAAGAGAGCGCT 7351  
 |||||  
 584 AspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrTh 600  
 |||||  
 7352 GACCAGGGGTGGTTGAAGGATGGTTCACACAGGTCTCCTTGATGACCAC 7401  
 |||||  
 600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 |||||  
 7402 CTGCTTTCTGCTCTGACGGGCCCTAGTAGTCTCTCTCTGTTACTTA 7451  
 |||||  
 617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633  
 |||||  
 7452 CAGTTGGGCTTGTCTTAATATAGATGTTTGTTCCTTTGTAGAGACGA 7501  
 |||||  
 634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLe 650  
 |||||  
 7502 GTGAGTGCAGTCCAGATCATGTACTTATAGGCAACAGTACCAAGGCTTCT 7551  
 |||||  
 650 uSerGlnGlyGluThrAspLeu 657  
 |||||  
 7552 GAGCCAAGGAGAACTGACCTC 7573

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV09700

seq\_documentation\_block:

ID AAV09700 standard; DNA; 8209 BP.

AC AAV09700;

XX 20-MAY-1998 (first entry)

XX Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.

XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;

XX virion core polypeptide; polymerase protein; envelope protein;

XX vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT misc\_feature 1..61

FT /tag= a

FT /note= "R LTR domain"

FT LTR 1..588

FT /tag= b

FT misc\_feature 62..143

FT /tag= c

FT /note= "U5 LTR domain"

FT CDS 588..2162

FT /tag= d

FT /product= GAG protein

FT /note= "viral core polypeptide"

FT CDS 2163..5747

FT /tag= e

FT /product= POL protein

FT /note= "polymerase peptide"

FT CDS 5620..7590

FT /tag= f

FT /product= ENV protein

FT /note= "envelope protein"

FT LTR 7591..8209

FT /tag= g

FT misc\_feature 7638..8106

FT /tag= h

FT /note= "U3 LTR domain"

FT protein\_bind 7643..7686

FT /tag= i

FT /note= "Binding site for transcription factors OCT-1, c-Myb, LyF-1, E47 and ETS-1"

FT protein\_bind 7719..7739

FT /tag= j

FT /note= "Binding site for transcription factors AP-4 and NF1"

FT protein\_bind 7744..7754

FT /tag= k

FT /note= "Binding site for transcription factor AP-1/TR"

FT protein\_bind 7773..7785

FT /tag= l

FT /note= "Binding site for transcription factor ETS-1/GATA"

FT protein\_bind 7793..7834

FT /tag= m

FT /note= "Binding site for transcription factors GATA, ETS-1, c-Myb, AP-1, GATA, E47"

FT protein\_bind 7841..7851

FT /tag= n

FT /note= "Binding site for transcription factor AP-1"

FT protein\_bind 7873..7883

FT /tag= o

FT /note= "Binding site for transcription factor AP-1"

FT protein\_bind 7885..7914

FT /tag= p

FT /note= "Binding site for transcription factors AP-1/GATA, c-Myb and AP-1"

FT protein\_bind 7928..7953

FT /tag= q

FT /note= "Binding site for transcription factors E47 and ETS-1"

FT CAAT\_signal 7964..7968

FT /tag= r

FT protein\_bind 7985..7997

FT /tag= s

FT /note= "Binding site for transcription factor GATA"

FT CAAT\_signal 8003..8007

FT /tag= t

FT protein\_bind 8024..8036

FT /tag= u

FT /note= "Binding site for transcription factor GATA"

FT CAAT\_signal 8042..8046

FT /tag= v

FT protein\_bind 8049..8060

FT /tag= w

FT /note= "Binding site for transcription factor AP-1/CREB"

FT TATA\_signal 8075..8090

FT /tag= x

FT misc\_feature 1..61

FT /tag= y

FT /note= "R LTR domain"

FT misc\_feature 8127..8135

FT /tag= z

FT /note= "PolyA downstream element"

FT polyA\_signal 8161..8166

FT /tag= aa

XX WO9740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997;

XX 97WO-GB01087.

XX 10-FEB-1997;

XX 97GB-0002668.

XX 19-APR-1996;

XX 96GB-0008164.

XX (IMUT-) IMUTRAN LTD.

XX (QONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;  
 XX WPI; 1997-535851/49.  
 DR P-PSDB; AAW39271; AAW39272; AAW39273.  
 XX  
 PT Polynucleotide encoding porcine retrovirus expression product -  
 PT useful to develop products for use in vaccines, diagnosis and  
 PT xeno-transplantation  
 XX  
 PS Claim 4; Fig 3; 69pp; English.  
 XX  
 CC This DNA sequence encodes the porcine retrovirus (PoEV) virion core  
 CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins and  
 CC also includes the Long Terminal Repeat (LTR). These proteins can be used  
 CC to develop viral vaccines, antisense nucleic acids, ribozymes and other  
 CC antiviral agents. They can also be used in xeno-transplantation  
 CC technology and as diagnostic tools.  
 XX  
 SQ Sequence 8209 BP; 2168 A; 2064 C; 2154 G; 1823 T; 0 other;

alignment\_scores:  
 Quality: 3444.50 Length: 658  
 Ratio: 5.275 Gaps: 2  
 Percent Similarity: 99.240 Percent Identity: 98.936

alignment\_block:  
 US-09-171-553B-10 x AAV09700 ..

Align seg 1/1 to: AAV09700 from: 1 to: 8209

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 5620 ATGCATCCACGTTAAAGCGCGCCACCTCCCGACTCGGGTGGAGAGCC 5669  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 5670 GAAAGAGCTGAGAAATCCCTTAAAGCTTCGGCTCCACGCTGGTTCCTTA 5719  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 5720 CTCCTAACAAATACCTCCCGAGCCAGTAGTAACGCGCTTATAGACAGCTCG 5769  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 5770 AACCCCATAGACCTTTATCCCTTACCTGGCTGATTTATGACCTGATAC 5819  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 5820 GGGTGTCACTGTAATAGCACCTCAGAGGTGTGCTCCTAGAGGCACCTGGT 5869  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 5870 GGCCGTGAACCTGCAATTCCTGCTCCGATGATTAACCCCGCTGTAAAGC 5919  
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 5920 ACACCTCCCAACCTAGTCCGTAGTATAGGGTCTTATGCTGCCAGGCAC 5969  
 117 rGluLysGluLysTyrCysGlyGlySerGlyGluSerPheCysArgArgT 134  
 5970 AGAGAAAGAGAAATACCTGTGGGGTTCCTGGGAATCCTTCTGTAGGAGAT 6019  
 134 rpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSerLeu 150  
 6020 GGAGCTCGGTCACTTCCAAAGATGGAGACTGGAAATGGCGGATCTCTCTC 6069  
 151 GlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrIly 167  
 6070 CAGGACCGGTAAATTCCTCTTTGTCATTCCTCGGCCCGGCAAGTACAA 6119  
 167 sMetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspT 184

|||||  
 6120 AATGATGAACACTATATAAAGATAAGAGCTCTCCCATCAGACTTAGATT 6169  
 184 yrLeuLysIleSerPheThrGluLys.GlyLysGlnGluAsnIleGlnLy 200  
 6170 ATCTAAAGATAAGTTTCACT...GAAAGGAAACAGGAAATATTTCAAA 6216  
 200 sTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyG 217  
 6217 GTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATTA.TATGGCGGG 6265  
 217 lyAlaGlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrGlu 233  
 6266 GAGCAGGGTCCACTTTAACCATTCGCTTAGGATAGAGAGGGGACAGAA 6315  
 234 ProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProPr 250  
 6316 CCCCCTGTGGCAATGGAGCCCGATAAGTACTGGCTGAACAGGGGCCCC 6365  
 250 oAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuA 267  
 6366 GGCCCTGGAGCCACCGCATAACTTGGGGTGGCCCAATTAACCTCGCTGC 6415  
 267 rgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIlePro 283  
 6416 GGCCTGACATAACACAGCGCCTAGCAACAGTACCACCTGGATTGATTCCT 6465  
 284 ThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGlnAr 300  
 6466 ACCAACCGCCTAGAAACTCCCGAGGTGTCTGTTAAGACAGGACAGAG 6515  
 300 gLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAsp 317  
 6516 ACTCTTCAGTCTCATCCAGGAGCTTTCGAAGCCATCAACTCCACGACC 6565  
 317 roAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProProTyr 333  
 6566 CTGATGCCACTTCTTCTGTGGCTTTGTCTATCTCTCAGGGCTCTCTTAT 6615  
 334 TyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisAr 350  
 6616 TATGAGGGGATGGCTAAAGAAAGAAATTTCAATGTGACCAAGAGCATAG 6665  
 350 gAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValS 367  
 6666 AAATCAATGTACATGGGGTCCGGAATAAGCTTACCTCCTCAGTGAATTT 6715  
 367 erGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeu 383  
 6716 CCGGGAAGGGGACATGCATAGGAAAGCTCCCGCATCCCAACACACCTT 6765  
 384 CysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeuVa 400  
 6766 TGCTATAGTACTGTGTTTATGAGCAGGCTCAGAAATTCAGTATTTAGT 6815  
 400 lProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCysV 417  
 6816 ACCTGGTTATAACAGGTGGTGGCATGCANTACTGGTTTAAACCCCTGTG 6865  
 417 alSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValGln 433  
 6866 TTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCAA 6915  
 434 IleValProArgValTyrTyrHisProGluGluValValLeuAspGluTy 450  
 6916 ATCGTCCCCCGAGTGTACTTACCATCTCTGAGGAAGTGGTCTTGATGAATA 6965  
 450 rAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuA 467  
 6966 TCAGTATCGGTATACCGGACCAAGAGAACCCGTATCCCTTACCTTAG 7015  
 467 laValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAla 483  
 |||||

7016 CTGTAATGCTCGATTAGGACGGCGTTGGCTAGGAAACAGGGACAGCT 7065  
 484 AlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuGlyGluLeuHi 500  
 |||||  
 7066 GCCTGATCACAGGACCACAGCAGCTAGAGAAGGACTTGGTGAGCTACA 7115  
 500 sAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAsnL 517  
 |||||  
 7116 TGGGGCCATCAGACAAGATCTCGAGCCTTAAAGGAGTCTGTTAGCAACC 7165  
 517 euGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArg 533  
 |||||  
 7166 TAGAAGAGTCCCTGACTCTTGTCTGAAGTGGTTCTACAGAACCGGAGG 7215  
 534 GlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLeu 550  
 |||||  
 7216 GGATTAGATCTGCTGTTCTTAAGAGAAGGTGGTTATGTGACGCTTAA 7265  
 550 sGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerM 567  
 |||||  
 7266 AGAAGAATGTTGCTTCTATGTATGATCCTCAGGAGCCATCAGAGACTCCA 7315  
 567 eAsnLysLeuArgLysLysLeuGluArgArgArgArgGluAla 583  
 |||||  
 7316 TGAACAAGCTTGAATAAAGTTAGAGAGCGTCGAGGGAAGAGAGGCT 7365  
 584 AspGlnGlyTyrPheGluGlyTrpPheAsnArgSerProTirpMetThr 600  
 |||||  
 7366 GACCAGGGTGGTTTCAAGGATGGTTCAACAGGTCTCCTTGGATGACCAC 7415  
 600 rLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeu 617  
 |||||  
 7416 CCTGCTTTCTGCTCTGACGGGCGCCCTAGTAGTCTGCTGCTGTTACTTA 7465  
 617 hrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArg 633  
 |||||  
 7466 CAGTGGGCTTGCCTTAATTAATAGTTTGTTCCTTGTAGAGAACA 7515  
 634 ValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLe 650  
 |||||  
 7516 GTGAGTGCAGTCCAGATCATGTTAGTCTTAGGCAACAGTACCAAGGCTTCT 7565  
 650 uSerGlnGlyGluThrAspLeu 657  
 |||||  
 7566 GAGCCAGGAGAACTGACCTC 7587

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAV82748

seq\_documentation\_block:

ID AAV82748 standard; DNA; 2462 BP.

XX AC AAV82748;

XX DT 25-FEB-1999 (first entry)

XX DE Pig endogenous retrovirus (PERV)-A envelope (env) gene region.

XX KW Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;  
 KW probe; primer; detection; retrovirus; human tissue; xenotransplant;  
 KW primary porcine tissue; human cell line; porcine cell line; ss.

XX OS Pig endogenous retrovirus.

XX FH Key Location/Qualifiers

FT CDS 211..2193

FT /\*tag= a

FT /product= envelope\_protein

XX PN W09853104-A2.

XX PD 26-NOV-1998.

XX PF 18-MAY-1998; 98WO-GB01428.

XX 16-MAY-1997; 97GB-0010154.  
 XX (MEDI-) MEDICAL RES COUNCIL.  
 XX Stoye JP, Weiss RA;  
 XX WPI; 1999-045324/04.  
 DR P-PSDB; AAW85453.  
 XX Newly isolated nucleic acid probe capable of hybridising to either  
 PT the PERV-A or PERV-B env gene - useful in the detection of  
 PT retroviruses, and their subtypes, in a sample of porcine/human  
 PT tissue  
 XX Claim 6; Page 20-21; 36pp; English.  
 XX The present sequence encodes a pig endogenous retrovirus (PERV)-A  
 CC envelope protein. PERV exists in two different subtypes, PERV-A and  
 CC PERV-B. The differences are reflected in sequence divergence in the  
 CC envelope genes. Probes and primers can be derived from the envelope  
 CC (env) genes of PERV-A and PERV-B. The probes and primers are used in  
 CC a method to detect retroviruses in a sample of porcine/human tissue,  
 CC particularly primary porcine tissue and human cell lines that have been  
 CC cultivated in the presence of a porcine cell line, or human tissue from  
 CC a patient with a xenotransplant. Subtype of PERV in a sample containing  
 CC one of the PERV env genes can also be determined.  
 XX Sequence 2462 BP; 715 A; 573 C; 575 G; 599 T; 0 other;

alignment\_scores:

Quality: 2631.50 Length: 679

Ratio: 4.468 Gaps: 9

Percent Similarity: 86.745 Percent Identity: 73.049

alignment\_block:

US-09-171-553B-10 x AAV82748 ..

Align seg 1/1 to: AAV82748 from: 1 to: 2462

1 MethisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 |||||  
 211 ATGCATCCACGTTAAGCGCGCCACCTCCCGATTCCGGGTGGAAGCC 260  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34  
 |||||  
 261 GAAAGAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTGTCTCTTA 310  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 |||||  
 311 CTCTGCAATAACTCTCAAGTTAATGTTAAACGCTTGTGGACAGCCCG 360  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||||  
 361 AACTCCCATATAACCCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 410  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84  
 |||||  
 411 AGGTATTAAATTAACAGCACTAAGGGGAGGCTCCCTTGGGGACCTGGT 460  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98  
 |||||  
 461 GGCCTGAATTATATGCTGCTTCGATCAGTATATCCCTGGTCTCATATGAC 510  
 99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115  
 |||||  
 511 CAGGCCACACCCCGATGACTCCGTCGTACGGGTTTACGGTTTTCGCC 560  
 115 oGly...ThrGluLysGluLysTyrCysGlyGlySerGlyGluSerPheC 131  
 |||||  
 561 AGGACCCCAATAATGAAGAAATATTGTGAAATCCCTCAGGATTTCTTTT 610

```

131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
|||||
611 GCAAGCAATGGAGCTGCATACCTCTTAATGATGGAAATGGAAATGGCCA 660
|||||
148 lIleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
|||||
661 GTCTCTCAGCAAGACAGAGTAAGTACTCTTTGTTTAAACAATCCTACCAG 710
|||||
162 .....GlyProGlyLysTyrLys.....MetMetL 170
|||||
711 TTATAATCAATTTAATTTATGGCCATGGGAGATGGAAAGATGGCAACAGC 760
|||||
170 ysLeuTyrLysAspLys.....SerCysSerProSerAsp 181
|||||
761 GGGTACAAAAGAGATGTACGAAATAAGCAATAAGCTGCTCATCTCGTTAGAC 810
|||||
182 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnII 198
|||||
811 CTAGATTACTTAAATAATAAGTTTCACCTGAAAAAGGAAAAACAAGAAATAT 860
|||||
198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrG 215
|||||
861 TCAAAAGTGGGTAATGGTATATCTTGGGAATAGTGTAC.....TATG 904
|||||
215 lyGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgIle 228
|||||
905 GAGGCTCTGGGAGAAAGAAAGGATCTGTTCTGACTATTCGCCCTCAGAAAT 954
|||||
229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245
|||||
955 GAAACTCAGATGGAACCTCCGGTGTCTATAGGACCAATAAAGGTTTGGC 1004
|||||
245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValProG 262
|||||
1005 CGAACAAAGGACCTCCCAATCCCAAGAACAG..... 1032
|||||
262 lnLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 278
|||||
1033 .....AGGCCATCTCCTAACCCCTCTGTATTACATACACACC 1068
|||||
279 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 295
|||||
1069 TCTGGATCAGTCCCACT.....GAGCCTAACATCACTAT 1103
|||||
295 lLysThrGlyClnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 312
|||||
1104 TAAACAGGGGGCAAGCTTTTACCCCTATCCAGGGAGCTTTTCAAGCTC 1153
|||||
312 leAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 328
|||||
1154 TTAACCTCCAGGACTCCAGAGGCTACCTCTTCTTGGCTTTGCTTAGCT 1203
|||||
329 SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnVa 345
|||||
1204 TCGGGCCCACTTACTATAGGGAATGGCTAGAGAGGAGGAATTCATATGT 1253
|||||
345 lThrLysLysHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuT 362
|||||
1254 GACAAAGGACATAGACCAATATGATAGGGATGCCAAATAAAGCTTA 1303
|||||
362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378
|||||
1304 CCCTTACTAGGTTTCTGGAAAGGACCTTCATAGGATGGTTCGCCCA 1353
|||||
379 SerHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGI 395
|||||
1354 TCCCAACCAACACCTTTCTAACCACTGAAGCCTTTAATCGAACCTCTGA 1403
|||||
395 uAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrG 412
|||||
1404 GAGTCAATATCTGTTACTGTTTATGACAGAGGTGGGCTGTAATAACTG 1453
|||||
412 lyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeu 428

```

```

1454 GATTAACCCCTTGTTTCCACCTTGTTTCAACCAAACTAAAGACTTT 1503
|||||
429 CysValMetValGlnIleValProArgValTyrHisProGluGluVa 445
|||||
1504 TCGGTTATGGTCCAAATTCGCCCGGGTGTACTACTATATCCGGAAGG 1553
|||||
445 ValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProV 462
|||||
1554 AGTCCTTGATGATATGACTATAGATATATCGGCCAAAGAGAGACCCA 1603
|||||
462 aIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
1604 TATCCCTGACACTAGCTGTAATGCTCGGATTGGGAGTGGCTCAGCGGTG 1653
|||||
479 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGI 495
|||||
1654 GGAACAGGAACGGCTGCCCTAATACAGGACCGCAACAGCTGGAGAAAG 1703
|||||
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
1704 ACTTAGTAACCTACATCGAATTGTACGGAAGATCTCCRAAGCCCTAGAHA 1753
|||||
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
1754 AATCTGTCACTAACCTGGAGGAATCCCTAACCTCTTATCTCAAGTGGTT 1803
|||||
529 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLe 545
|||||
1804 CTACAGAACAGAGGGGGTTAGATCTGTTATTTCTAAAAGAGAGGGGTT 1853
|||||
545 uCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyA 562
|||||
1854 ATGTGTAGCCTTAAAGAGGAATGCTGCTCTATGTAGATCACTCAGGAG 1903
|||||
562 laIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArg 578
|||||
1904 CCATCAGAGACTCCATGACCAAGCTTAGAAAAAGGTTAGAGAGGCTCGA 1953
|||||
579 ArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSe 595
|||||
1954 AGGGAAGAGAGAGGCTGACAGGGGGTGGTTTGAAGGATGGTTTCAACAGGTC 2003
|||||
595 rProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValL 612
|||||
2004 TCTTGGATGACCACTCTCTTCTGCTCTGACGGGGCCCCCTAGTAGTCC 2053
|||||
612 euLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 628
|||||
2054 TGCTCTGTTACTTACAGTTGGCCTTGCTTAATTAATAGGTTTGTGGCC 2103
|||||
629 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGI 645
|||||
2104 TTTGTTAGAGAACGAGTAGTGCAGTCCAGATCATGGTACTTAGGCAACA 2153
|||||
645 nTyrGlnGlyLeuLeuSerGlnGlyGluThrAspLeu 657
|||||
2154 GTACCAAGGCCCTCTGAGGCCAAGGAGAACTGACCTC 2190
|||||

```

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67022

seq\_documentation\_block:

ID AAC67022 standard; DNA; 4918 BP.

XX AAC67022;

XX AC (first entry)

XX 27-MAR-2001

XX PERV env protein coding sequence SEQ ID NO: 22.

XX Xenotransplantation; infectious agent; vaccine; ds.

XX Porcine endogenous retrovirus.

XX WO2000071726-A1.  
 PN 30-NOV-2000.  
 PD 24-MAY-2000; 2000WO-US14296.  
 XX 24-MAY-1999; 99US-0135631.  
 XX (MAYO-) MAYO MEDICAL VENTURES.  
 PA Federspiel MJ;  
 PI WPI; 2001-032041/04.  
 DR Inhibiting or preventing infectious agent transmission in mammalian  
 XX transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -  
 XX  
 XX Claim 16; Page 109-111; 144pp; English.  
 PS The present invention provides a method to prevent the transmission of  
 XX infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.  
 XX  
 XX Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;

## alignment\_scores:

Quality: 2549.50 Length: 669  
 Ratio: 4.411 Gaps: 9  
 Percent Similarity: 86.398 Percent Identity: 71.001

## alignment\_block:

US-09-171-553B-10 x AAC67022 ..  
 Align seg 1/1 to: AAC67022 from: 1 to: 4918  
 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 |||||  
 2187 ATGCATCCACGTTAAAGCGCGCCACCTCCCGATTCCGGGTGGAAAGCC 2236  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeu 34  
 |||||  
 2237 GAAAGACTGAAATCCCTTAAAGCTTGCCTCCATCCGCTGGTTCCTTA 2286  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 |||||  
 2287 CTCGTCAATAACTCTCAAGTTAATGGTAACGCCCTTGGGACAGCCCG 2336  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||||  
 2337 AACTCCCAATAAACCTTATCTCTCACCTGGTTACTTACTGACTCCGGTAC 2386  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84  
 |||||  
 2387 AGGTATTAAATTAACAGCACTCAAGGGGAGGCTCCCTTGGGGACCTGGT 2436  
 84 rProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal. .... 98  
 |||||  
 2437 GGCCTGAATTATATCTGCTTCGATCAGTAATCCCTGGTCTCAATGAC 2486  
 99 LysSerThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysPr 115  
 |||||  
 2487 CAGGCCACACCCCGGATGTAACCGTCTTACGGGTTTACGTTTGGCC 2536  
 115 oGly...ThrCluLysGluLysTrpCysGlyGlySerGlyGluSerPheC 131  
 |||||  
 2537 AGGACCCCAATAATGAAGAATATTGTGAAATCCCTCAGGATTTCTTTT 2586

131 ySArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147  
 |||||  
 2587 GCAAGCAATGGAGCTGCTAACTTCTAATGATGGAAATGGAAATGGCCA 2636  
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer. .... 161  
 |||||  
 2637 GTCTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTTAACAATCCTACCAG 2686  
 162 .....GlyProGlyLysTrpLys.....MetMetL 170  
 |||||  
 2687 TTATAATCAATTTAATTTATGGCCATGGAGATGGAAGATGGCAACAGC 2736  
 170 ySLeuTrpLysAspLys.....SerCysSerProSerAsp 181  
 |||||  
 2737 GGGTACAAAAGATGTACGAAATAAGCAATAAGCTGTCTATTCGTTAGAG 2786  
 182 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIl 198  
 |||||  
 2787 CTAGATTACTTAAATAAGTTTCACTGAAAAAGGAAAAACAAGAAATAT 2836  
 198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTrpLysTrp 215  
 |||||  
 2837 TCAAAAGTGGTAATGGTATGCTTTGGGAATAGTGATC.....TATG 2880  
 215 yGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgIle 228  
 |||||  
 2881 GAGCTCTGGGAGAAAGAAAGATCTGTTCTGACTATTCGCTCAGAATA 2930  
 229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245  
 |||||  
 2931 GAACTCAGATGGAACTCCGGTTGCTATAGGACCAATAAGGGTTGGC 2980  
 245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValPro 262  
 |||||  
 2981 CGAACAGGACCTCCCAATCCAGAACAG..... 3008  
 262 lnLeuThrSerLeuArgProAspIleThrClnProProSerAsnSerThr 278  
 |||||  
 3009 .....AGCCATCTCTTAACCCCTCTGATTACATAACAAC 3044  
 279 ThrGlyLeuIleProThrAsnThrProArgAsnSerProGlyValProVa 295  
 |||||  
 3045 TCTGGTCACTCCCACT.....GAGCCTAACATCACTAT 3079  
 295 lLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI 312  
 |||||  
 3080 TAAACAGGGCGGAACTTTTAACTCATCCAGGAGCTTTTCAAGCTC 3129  
 312 leAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer 328  
 |||||  
 3130 TTAACCTCCAGACTCCAGAGGCTTCTTTGTTGGCTTTGCTTAGCT 3179  
 329 SerGlyProProTrpTrpGluGlyMetAlaLysGluArgLysPheAsnVa 345  
 |||||  
 3180 TCGGGCCCACTTACTATCAGGGAATGGCTAGAGGAGGGAATTCATGT 3229  
 345 lThrLysGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeu 362  
 |||||  
 3230 GACAAAGAACATAGACCAATGTACATGGGATCCCAATAAGCTTA 3279  
 362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378  
 |||||  
 3280 CCTTACTGAGGTTCTGGAAGGACCTGCTAGGATGGTTCCTCCCA 3329  
 379 SerHisGlnHisLeuCysTrpSerThrValValTrpGluGlnAlaSerGl 395  
 |||||  
 3330 TCCCAACACACCTTGTAAACACACCTGAAGCCTTAACTCAACCTCGA 3379  
 395 uAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrG 412  
 |||||  
 3380 GAGTCAGTATCTGGTACCTGTTATGACAGGTGGTGGCATGTAATACTG 3429  
 412 lLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeu 428



```

|||||
3430 GATTAACCCCTGTGTTCACCTTGGTTTCAACCAACTAAAGACTTT 3479
|||||
429 CysValMetValGlnIleValProArgValTyrHisProGluGluVa 445
|||||
3480 TGCCTTATGTCCTAAATGTCCCGGGGTGTACTTACTATCCGAAAAAGC 3529
|||||
445 lValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProV 462
|||||
3530 ACTCCTTGATGAATATGACATATAGATATATCGGCCAAAGAGAGCCCA 3579
|||||
462 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
3580 TATCCCTGACACTAGCTAGCTGTAATGCTCGGATTGGAGTGGCTGCAGGCGTG 3629
|||||
479 GlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysG 495
|||||
3630 GGAACAGGAAGCGGTGCCCTAATCACAGAGACCCGCAACAGCTGGAGAAGG 3679
|||||
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
3680 ACTTAGTAACCTACATCGAATTGTAACGGAAGATCTCCAAGCCCTAGAAA 3729
|||||
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
3730 AATCTGTCACTAGTAACCTGAGGAATCCCTAACCTCTTATCTGAAGTGGTT 3779
|||||
529 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyLe 545
|||||
3780 CTACAGAACAGAGGGGTTAGATCTCTTTATTTCTAAAGAGGAGGATT 3829
|||||
545 uCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyA 562
|||||
3830 ATGTGTAGCCTTGAAGAGGAATGCTGTTTTATGTGGATCATTCAGGGG 3879
|||||
562 laIleArgAspSerMetAsnLysLeuArgLysLysLeuGluArgArgArg 578
|||||
3880 CCATCAGAGACTCCATCAACAAGCTTAGAGAAAGGTTGGAGAAAGCGTCA 3929
|||||
579 ArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSe 595
|||||
3930 AGGAAAAGAAACACTACTCAAGGTTGGTTTTCAGGATGGTTCAACAGGTC 3979
|||||
595 rProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVal 612
|||||
3980 TCCTTGGTGGTACCTTCTTCTGCTTTAACAGGACCCCTTATATAGTCC 4029
|||||
612 euLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla 628
|||||
4030 TCCTCTCTGTACTCACAGTTGGGCCATGTATTATTAAACAAGTTAATTGCC 4079
|||||
629 PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnG 645
|||||
4080 TTCATTAGAACGAATTAAGTCAGTCCAGATCATGTGTTACTTACAGAAC 4129
|||||
645 nTyrGln 647
|||||
4130 GTACCAA 4136

```

seq\_name: /Sids2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67018

seq\_documentation\_block:

ID AAC67018 standard; DNA; 1980 BP.

XX AAC67018;

XX AC AAC67018;

XX 27-MAR-2001 (first entry)

XX PERV-1-15 env protein coding sequence SEQ ID NO: 18.

XX KW Xenotransplantation; infectious agent; vaccine; ds.

XX OS Porcine endogenous retrovirus.

```

XX WO200071726-A1.
XX 30-NOV-2000.
XX 24-MAY-2000; 2000WO-US14296.
XX 24-MAY-1999; 99US-0135631.
XX (MAYO-) MAYO MEDICAL VENTURES.
XX Federspiel MJ;
XX WPI; 2001-032041/04.
XX Inhibiting or preventing infectious agent transmission in mammalian
XX transplant recipients, by introducing recombinant DNA comprising DNA
XX encoding extracellular proteins of the agent into donor cells, such as
XX swine cells -
XX Claim 16; Page 100-101; 144pp; English.
XX The present invention provides a method to prevent the transmission of
XX infectious agents during xenotransplantation. This involves introducing
XX to donor swine cells a recombinant DNA encoding a peptide fragment from
XX the infectious agent, and then introducing these cells into the
XX transplant recipient.
XX Sequence 1980 BP; 582 A; 452 C; 447 G; 499 T; 0 other;

```

alignment\_scores:

Quality: 2546.50

Length: 667

Ratio: 4.413

Gaps: 8

Percent Similarity: 86.507

Percent Identity: 70.915

alignment\_block:

US-09-171-553B-10 x AAC67018

Align seg 1/1 to: AAC67018 from: 1 to: 1980

```

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
1 ATGCATCCACCGTTAAGCGCGCCACCTCCCGATTCCGGGTGGAAGCC 50
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
51 GAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTTCCTTA 100
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
101 CTCTGTCAATTAACCTCTCAAGTTAATGGTAAACGCTTGTGCACAGCCG 150
51 AsnProHisArgProLeuSerProThrTrpLeuIleAlaAspProAspTh 67
151 AACTCCCATAAACCCCTTATCTCTACCTGGTTACTTACTGACTCCGGTAC 200
67 rglyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
201 AGTATTATATTAACAGCACTCAAGGGAGGCTCCCTTGGGGACCTGCT 250
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaVal..... 98
251 GGCTTGATTAATATGCTCGCTTCGATCATGTAATCCCTGGTCTCAATGAC 300
99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115
301 CAGGCCACACCCCGGATGATGCTACTCGTGTCTTACGGGTTTACGTTGCC 350
115 oGly...ThrGluLysGluLysTyrCysGlySerGlyGlyGluSerPheC 131
351 AGGGCCCCCAATAATGAAGAATATTGTGAAATCTCAGGATTCTTTT 400

```

```

131 ySArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147
|||||
401 GCAAGCAATGGAGCTCGTAACCTCTAATGATGGGAATGGAAATGGCCA 450
|||||
148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161
|||||
451 GTCTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTAAACAATCTACCAG 500
|||||
162 .....GlyProGlyLysTyrLys.....MetMetL 170
|||||
501 TTATAATCAATTTAATATATGGCCATGGGAGATGGAAAGATTGGCAACAGC 550
|||||
170 ySLeuTyrLysAspLys.....SerCysSerProSerAsp 181
|||||
551 GGGTACAAAAGATGTACCAAAATAGCAAAATAGCTGCATTCGTTAGAC 600
|||||
182 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnI 198
|||||
601 CTAGATTACTTAAAAATAAGTTTCACTGAAAAAGGAAACAAAGAAATAT 650
|||||
198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrG 215
|||||
651 TCAAAAGTGGGTAAATGGTATCTCTGGGGAATAGTGTACTATAGAGCT 700
|||||
215 lyGly...GlyAlaGlySerThrLeuThrIleArgLeuArgIleGluThr 230
|||||
701 CTGGGAGAAAGAAAGATCTGTCTGACTATTCGCCCTCAGAAATAGAACT 750
|||||
231 GlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGlu 247
|||||
751 CAGATGGAACCTCCGGTTCCTATAGACCAATAAGGGTTTGGCGGAACA 800
|||||
247 nGlyProProAlaLeuGluProProHisAsnLeuProValProGlnLeu 264
|||||
801 AGGACCTCCAATCCAAGAACAG..... 822
|||||
264 hrSerLeuArgProAspIleThrGlnProProSerAsnSerThrGly 280
|||||
823 .....AGGCATCTCTTAACCCCTCTGATTACAATACACCTCTCGA 864
|||||
281 LeuIleProThrAsnThrProArgAsnSerProGlyValProValLysTh 297
|||||
865 TCAGTCCGCCACT.....GAGCCTAACATCACTATTAAAAAC 899
|||||
297 rGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnS 314
|||||
900 AGGGCGAAACTTTTAACTCATCCAGGAGCTTTTCAAGCTCTTAACT 949
|||||
314 erThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGly 330
|||||
950 CCACGACTCCAGAGCTACCTCTCTCTGCTTGGCTTGGCTTACGCTCGGC 999
|||||
331 ProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLy 347
|||||
1000 CCACCTTACTATGAGGGAATGGCTAGAGGAGGAAATTCAAATGTGACAAA 1049
|||||
347 sGluHisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeu 364
|||||
1050 GGAACATAGACCACTTATGATGGGATGCCAAATAAGCTTACCCCTTA 1099
|||||
364 hrGluValSerGlyLysGlyThrCysIleGlyLysAlaProProSerHis 380
|||||
1100 CTGAGGTTTCTGGAAAAGGCACCTGCATAGGATGGTTCCTCCCATCCAC 1149
|||||
381 GlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerGluAsnG 397
|||||
1150 CAACACCTTGTGACCAACACCTGAAGCCTTTAATCGAACCTCTGAGAGTCA 1199
|||||
397 nTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeu 414
|||||
1200 GTAICTGTACTGGTATGACAGTGTGGGCATGTAATCTCGATTA 1249
|||||
414 hrProCysValSerThrSerValPheAsnGlnSerLysAspLeuCysVal 430
|||||

```

```

1250 CCCCTTGTTTCCACCTTGTTTTCACCAAACTAAAGACTTTTGCGTT 1299
|||||
431 MetValGlnIleValProArgValTyrTyrHisProGluGluValValLe 447
|||||
1300 ATGTCACAAATTTGCCCGGGTGTACTACTATCCCGAAAAAGACAGTCT 1349
|||||
447 uAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerL 464
|||||
1350 TGATGAATATGACTATAGATATAATCGCCAAAAAGAGAGCCCATATCC 1399
|||||
464 euThrLeuAlaValMetLeuGlyThrAlaValAlaGlyValGlyThr 480
|||||
1400 TGACACTAGCTGTAATGCTCGATGGAGTGGCTGCAGCGGTGGGAACA 1449
|||||
481 GlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuG 497
|||||
1450 GGAAGCGCTGCCCTAATCACAGGACCCACACAGCTGGAGAAAGACTTAG 1499
|||||
497 yGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerV 514
|||||
1500 TAACCTACATCGAATTTGTAACGGAAAAATCTCCAAGCCCTAGAAAAATCTG 1549
|||||
514 alSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeuGln 530
|||||
1550 TCAGTAACCTGGAGGAATCCCTAACCTCCTTATCTGAAGTGGTCTACAG 1599
|||||
531 AsnArgArgGlyLeuAspLeuLeuPheLeuArgGlyGlyLeuCysAl 547
|||||
1600 AACAGAAGGGGTAGATCTGTTATTTCTAAAGAGGAGGATATGTGT 1649
|||||
547 alAlaLeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleA 564
|||||
1650 AGCCTTAAGGAGGAATGCTGTTTTTATGTGGATCATTCAGGGGCCATCA 1699
|||||
564 rGaspSerMetAsnLysLeuArgLysLysLeuGluArgArgArgGlu 580
|||||
1700 GAGACTCCATGAACAAGCTTAGAGAAAGGTGGAGAGCGCTCGAAGGGAA 1749
|||||
581 ArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTr 597
|||||
1750 AAGGAAACTACTCAAGGGTGTGTTTGAAGGATGTTCAACAGGCTCCTTC 1799
|||||
597 pMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeu 614
|||||
1800 GTTGGCTACCTCTTCTCTTAAACAGGACCTTAAATAGTCTCCTCC 1849
|||||
614 euLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheVal 630
|||||
1850 TGTACTCAGAGTGGCCATGTATTATTAAACAAGTTAATTCCTTCATT 1899
|||||
631 ArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrG 647
|||||
1900 AGAGAAGCAATAAGTGCAGTCCAGATCATGTTACTTGTAGACAACAGTACCA 1949
|||||
647 n 647
1950 A 1950
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77726
seq_documentation_block:
ID AAF77726 standard; DNA; 7333 BP.
XX
AC AAF77726;
XX
DT 23-MAY-2001 (first entry)
XX
DE Defective retroviral genome isolated from PK-15 cell line.
XX
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;
KW ss.
XX

```





## alignment\_scores:

Quality: 2537.50 Length: 658  
Ratio: 4.452 Gaps: 9  
Percent Similarity: 86.626 Percent Identity: 72.796

## alignment\_block:

US-09-171-553B-10 x AAT74883 ..

Align seg 1/1 to: AAT74883 from: 1 to: 7393

22 lleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleTh 38  
4806 ATCCCTTAAGCTTCGCCCTCCATCGCGTGGTTCCTTACTCTGTCAATAAC 4855  
38 rProGlnAlaSerSerLysArgLeuIleAspSerSerAsnProHisArgp 55  
4856 TCCTCAAGTTAATGGTAACGCCCTTGTGGACAGCCCGAACCTCCCATAAAC 4905  
55 toLeuSerProThrTrpLeuIleIleAspProAspThrGlyValThrVal 71  
4906 CCTTATCTCTACCTGGTTACTTACTGACTCCGGTACAGGTATTAATATT 4955  
72 AsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeuHi 88  
4956 AACAGCACTCAAGGGGAGCTCCCTTGGGACCTGGTGGCCTGAATTATA 5005  
88 spheCysLeuArgLeuIleAsnProAlaVal.....LysSerThrProp 103  
5006 TGCTGCCTTCGATCAGTAATCCCTGGTCTCAATCAGCAGCCACACCCC 5055  
103 roAsnLeuValArgSerTyrGlyPheTyrCysCysProGly...ThrGlu 118  
5056 CCGATGACTCCGGTCTACGGGTTCACGTTGCCCGGACCCCAAT 5105  
119 LysGluLysTyrCysGlySerGlyLeuSerPheCysArgArgTrpSe 135  
5106 AATGAAGAATATTGGGAATCCTCAGGATTTCTTTTGCAGCAATGGAG 5155  
135 rCysValThrSerAsnAspGlyAspTrpLysTrpProLysLeuGlnA 152  
5156 CTGCATAACTTCTAATGATGGAAATGGAAATGGCCAGTCTCTCAGCAAG 5205  
152 spArgValLysPheSerPheValAsnSer..... 161  
5206 ACAGATGAAGTTACTCTTTGTTAACAACTCTACAGTTATATCAATTT 5255  
162 .....GlyProGlyLysTyrLys.....MetMetLysLeuTyrLysAs 174  
5256 AATTATGGCCATGGGAGATGGAAAGATTGGCAACAGCGGGGTACAAAAGA 5305  
174 pLys.....SerCysSerProSerAspLeuAspTyrLeuL 186  
5306 TGTACGAAATAAGCAAAATAAGCTCTCATTCGTTAGACCTAGATTACTTAA 5355  
186 ysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrpIle 202  
5356 AAATAGTGTTCACGTGAAGAGGAAACAAAGAAATATTCAAAAGTGGGTA 5405  
203 AsnGlyMetSerTrpGlyIleValPheTyrLysTyrGlyGlyAla... 218  
5406 AATGTTATCTCTGGGAATAGTGAC.....TATGGAGGCTCTGGGAG 5449  
219 .....GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThrG 233  
5450 AAGAAAGAGGATCTCTTCTGACTATTCCCTCAGAAATAGAAACTCAGATGG 5499  
233 luProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyPro 249  
5500 AACCTCGGTTGCTATAGGACCAAAATAAGGTTTGGCCGAAACAGGACCT 5549  
250 ProAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLe 266

5550 CCAATCCAAGACAG..... 5564  
266 uArgProAspIleThrGlnProProSerAsnSerThrThrGlyLeuIleP 283  
5565 .AGCCATCTCCTAAACCCCTCTGATTACAATAACAACCTCTGGATCAGTCC 5613  
283 roThrAsnThrProArgAsnSerProGlyValProValLysThrGlyGln 299  
5614 CCACT.....GAGCCTAACCATCACTATTAAACAGGGGGG 5648  
300 ArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAsnSerThrAs 316  
5649 AAACCTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGCAG 5698  
316 pProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 333  
5699 TCCAGAGGCTACCTCTCTTGTGGTCTTGTAGCTTCGGGCCACCTT 5748  
333 yrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHis 349  
5749 ACTATGAGGAATGGCTAGAGGAGGAATTCATGTGACAAAGGAACAT 5798  
350 ArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluVa 366  
5799 AGAGACCAATGTATACATGGGGATCCCAAAATAAGCTTACCTTACTGAGGT 5848  
366 lSerGlyLysGlyThrCysIleGlyLysAlaProProSerHisGlnHisL 383  
5849 TTCTGAAAAAGGCACCTGCTAGGATGGTTCCCCATCCCAACACACC 5898  
383 euCysTyrSerThrValValTyrGluGlnAlaSerGluAsnGlnTyrLeu 399  
5899 TTTGTAACCACTGAAGCCTTTAATCGAACCTCTCAGAGTCAATATCTG 5948  
400 ValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGlyLeuThrProCy 416  
5949 GTACCTGGTTATGACAGGTGGGGCATGTAATACCTGGATTACCCCTTG 5998  
416 sValSerThrSerValPheAsnGlnSerLysAspLeuCysValMetValG 433  
5999 TGTGTTCCACCTTGGTTTTCACCAACAACTAAAGACTTTTGGTTATGGTCC 6048  
433 lntIleValProArgValTyrTyrHisProGluGluValValLeuAspGlu 449  
6049 AAATGTCCTCCCGGGTGTACTACTATCCGAAAAAGACGCTCCTTGATGA 6098  
450 TyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLe 466  
6099 TATGACTATAGATATTAATCGGCCAAAAAGAGACCCATATCCTTGACACT 6148  
466 uAlaValMetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThra 483  
6149 AGCTGTAATGCTCGGATTGGGAGTGGCTCAGCGGTGGGAACAGAACGG 6198  
483 laAlaLeuIleThrGlyProGlnGlnGluLysGlyLeuGlyGluLeu 499  
6199 CTGCCCTTAATCAGGACCGCAACAGCTGGAGAAAGGACTTAGTAACCTTA 6248  
500 HisAlaAlaMetThrGluAspLeuArgAlaLeuLysGluSerValSerAs 516  
6249 CATCGAATTGTAAACGGAAGATCTCCAGGCCCTAGAAAAATCTGTAGTAA 6298  
516 nLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgA 533  
6299 CCTGGAGGAATCCCTAACCTCTTATCTGAAGTGGTCTTACAGAACAGAA 6348  
533 rgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeu 549  
6349 GGGGGTGTAGATCTGTATTCTTAAAGAAAGAGGAGGTATGTGTAGCCTTA 6398  
550 LysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSe 566

6399 AAGAGGAATGCTGCTTCTATCTAGATCACTCAGGAGCCATCAGAGACTC 6448  
 566 rMetAsnLysLeuArgLysLysLeuGluArgArgArgArgGluArgGluA 583  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6449 CATGAGCAAGCTTAGAGAAAGCTTAGAGAGGCGTCGAAGGAAAGAGAGG 6498  
 583 laaspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThr 599  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6499 CTGACAGAGGGTGGTTGAAGATGGTTCAACAGGTCTCCTTGGATGACC 6548  
 600 ThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuValLeuLeuLe 616  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6549 ACCCTGCTTCTGCTCTGACGGGGCCCTAGTAGCTCTGCTCTTACT 6598  
 616 uThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluA 633  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6599 TACAGTTGGCCTTGCTTAATTAATAGTTTGTGCTTTGTAGAGAAC 6648  
 633 rgValSerAlaValGlnIleMetValLeuArgGlnGlnTrpGlnGlyLeu 649  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6649 GAGTGAGTGCACTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTT 6698  
 650 LeuSerGlnGlyGluThrAspLeu 657  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6699 CTGAGCCAGGAGAACTGACCTC 6722

seq\_name: /SIP02/gcgdata/geneseq/geneseq/NA2001.DAT: AAC67023

seq\_documentation\_block:

ID AAC67023 standard; DNA; 7873 BP.

AC AAC67023;

DT 27-MAR-2001 (first entry)

DE PERV env protein coding sequence SEQ ID NO: 23.

KW Xenotransplantation; infectious agent; vaccine; ds.

OS Porcine endogenous retrovirus.

PN WO200071726-A1.

PD 30-NOV-2000.

PF 24-MAY-2000; 2000WO-US14296.

PR 24-MAY-1999; 99US-0135631.

PA (MAYO-) MAYO MEDICAL VENTURES.

PI Federspiel MJ;

DR WPI; 2001-032041/04.

PT Inhibiting or preventing infectious agent transmission in mammalian.  
 PT transplant recipients, by introducing recombinant DNA comprising DNA  
 PT encoding extracellular proteins of the agent into donor cells, such as  
 PT swine cells -

PS Claim 16; Page 112-115; 144pp; English.

CC The present invention provides a method to prevent the transmission of  
 CC infectious agents during xenotransplantation. This involves introducing  
 CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
 CC the infectious agent, and then introducing these cells into the  
 CC transplant recipient.

SQ Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

alignment\_scores:

Quality: 2514.50 Length: 703

Ratio: 4.358 Gaps: 10  
 Percent Similarity: 82.077 Percent Identity: 67.425  
 alignment\_block:  
 US-09-171-553B-10 x AAC67023 ..  
 Align seg 1/1 to: AAC67023 from: 1 to: 7873  
 1 MethHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5012 ATGCATCCACGTTAAGCGCGCGCACCTCCCGATTCCGGGTGGAAGCC 5061  
 17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5062 GAAAGAACTGAAATCCCTTAAGCTTCGCCTCCATCGCTGGTTCCTTA 5111  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5112 CTCGTCAATAACCCCTCAAGTTAAGTTAAACCCCTTGTGGACAGCCCG 5161  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5162 AACTCCCATAAACCCCTTATCTCTACCTTGGTTACTTACTCACTCCGCTAC 5211  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5212 AGGTATTAAATTAACAGACACTCAAGGGGAGGCTCCCTTGGGACCTGGT 5261  
 84 rpProGluLeuHisPheCysLysArgLeuIleAsnProAlaVal..... 98  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5262 GGCTGGAATTATAGTCTGCTTCGATCAGTAACTCCCTGGTCTCAATGAC 5311  
 99 LysSerThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysPr 115  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5312 CAGGCCACACCCCGATGTAACCGTCTTACGGGTTCCTACGTTTGGCC 5361  
 115 oGly...ThrGluLysGluLysTrpCysGlyGlySerGlyGluSerPheC 131  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5362 AGGACCCCAATAATGAAGAATATTGGAAATCCTCAGGATTTCTTTT 5411  
 131 ysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro 147  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5412 GCAAGCAATGGAGCTGCTAACTTCTAATGATGGAAATGGAAATGGCCA 5461  
 148 IleSerLeuGlnAspArgValLysPheSerPheValAsnSer..... 161  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5462 GTCTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTAAACAATCCTACCAG 5511  
 162 .....GlyProGlyLysTyrLys.....MetMetL 170  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5512 TTATAATCAATTAATTATGGCCATGGGAGATGGAAAGATTGGCAACAGC 5561  
 170 ysLeuTyrLysAspLys.....SerCysSerProSerAsp 181  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5562 GGGTACAAAAGATGTAGCAAAATAAGCAATAAGCTGTCTCATTCGTTAGAC 5611  
 182 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIl 198  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5612 CTAGATTACTTAAATAATAGTTTCACTGAAAAAGGAAACAAAGAAATAT 5661  
 198 eGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTrpG 215  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5662 TCAAAAGTGGTAATAGTATGCTTGGGAATAGTAGTAC.....TATG 5705  
 215 lyGlyGlyAla.....GlySerThrLeuThrIleArgLeuArgIle 228  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5706 GAGCTCTGGGAGAAAGAAAGGATCTGTCTGACTATTGCGCTCAGAATA 5755  
 229 GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAl 245  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 5756 GAAACTCAGATGGAAACCTCCGGTGTATAGGACCAATAAAGGGTTTGGC 5805  
 245 aGluGlnGlyProProAlaLeuGluProProHisAsnLeuProValProG 262

```

|||||
5806 CGAACAGGACCTCAATCCAGAACAG..... 5833
262 InLeuThrSerLeuArgProAspIleThrGlnProProSerAsnSerThr 278
|||||
5834 .....AGGCCATCTCTTAACCCCTCTGATTACATACACACC 5869
279 ThrGlyLeuLeuProThrAsnThrProArgAsnSerProGlyValProVa 295
|||||
5870 TCTGGATCAGTCCGCCACT.....GAGCCTAACACTACTAT 5904
295 IlyThrGlyGlnArgLeuPheSerLeuLeuGlnGlnAlaPheGlnAlaI 312
|||||
5905 TAAACAGGGCGAAACTTTTAAACCTCATCCAGGGAGCTTTTCAAGCTC 5954
312 LeAsnSerThrAspProAspAlaThrSerSerCysTyrLeuCysLeuSer 328
|||||
5955 TTAACCTCACGACTCCAGAGCTACCTCTCTTCTGCTTGGCTTGTAGCT 6004
329 SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnVa 345
|||||
6005 TCGGGCCACCTTACTACTATGAGGAATGGCTAGAGGAGGAAATTCATGT 6054
345 lThrLysGluHisArgAsnGlnCysThrTyrGlySerArgAsnLysLeuT 362
|||||
6055 GACAAAGGACATAGAGACCAATATCATGGGATCCCAAAATAAGCTTA 6104
362 hrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysAlaProPro 378
|||||
6105 CCCTTACTAGGTTCTGGAAAGGACCTGCATAGGGATGGTTCCTCCCA 6154
379 SerHisGlnHisLeuCysTyrSerThrValValTyrGluGlnAlaSerG 395
|||||
6155 TCCCAACCAACACCTTTGTAACACACATGAAGCCCTTTAATCGAACCTCTGA 6204
395 uAsnGlnTyrLeuValProGlyTyrAsnArgTyrTrpAlaCysAsnThrG 412
|||||
6205 GAGTCAGTATCTGGTACTCTGTTATGACAGAGTGGTGGCATGTAATACTG 6254
412 lyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeu 428
|||||
6255 GATTAACCCCTTGTTTCCACCTTGGTTTCAACCAACCAATAAGACTTT 6304
429 CysValMetValGlnIleValProArgValTyrTyrHisProGluGluVa 445
|||||
6305 TGCCTTATGGTCCAAATGTCCCGGGTGTACTACTATCCCGAAAAAGC 6354
445 lValLeuAspGluTyrAspTyrArgTyrAsnArgProLysArgGluProV 462
|||||
6355 AGTCCCTTGATCAATATGACTATATAGATATAATCGCCCAAAAGAGGCCCA 6404
462 alSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlaValGlyVal 478
|||||
6405 TATCCCTGCACACTAGCTGTAATGCTCGGATGGAGTGGCTGCAGCGGTG 6454
479 GlyThrGlyThrAlaAlaLeuLeuThrGlyProGlnGlnLeuGlyLysG 495
|||||
6455 GGAACAGGAACGGCTGCCCTTAATCACAGGACCGCAACAGCTGGAGAAAG 6504
495 yLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysG 512
|||||
6505 ACTTAGTAACCTACATCGAATTGTAACGGAAGATCTCCAAAGCCCTAGAA 6554
512 luSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal 528
|||||
6555 AATCTGTCAGTACCTGGAGGAAATCCCTAACCTCTTATCTGAAGTGGTT 6604
529 LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyLe 545
|||||
6605 CTACAGAACAGAGGGGTTAGATCTGTTATTTCTAAAGAGAGGAGATT 6654
545 uCysAlaAlaLeuLysGluCysCysPheTyrValAspHisSerGlyA 562
|||||

```

```

6655 ATGTGTAGCCTTGAAGGAGGAATGCTGTTTTTATGTGGATCATTCAGGGG 6704
562 laIleArgaspSerMetAsnLys..... 569
|||||
6705 CCATCAGAGACTCCATGACACAGACTTAGAGAAAGGACTGGAGAGCCCGG 6754
569 ..... 569
6755 CGGGTCTCGAACAAACCCAGACAGGTTCCTGTTTCAATTAAAGAACTGTC 6804
570 .....LeuArgLysLysLeuGluArgArgA 578
|||||
6805 GAAGTAACCGCTGAGCTAAAGCCAGCTTAGAGAAAGGTTGGAGAGCGTC 6854
578 rArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArg 594
|||||
6855 GAAGGAAAGAGAACTACTCAAGAGTGGTTTGAGGGATGGTTCAACAGG 6904
595 SerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValVa 611
|||||
6905 TCTCCTTGGTGGCTACCTACTTCTGCTTTAACAGGACCTTAATAGT 6954
611 lLeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgPheValA 628
|||||
6955 CTCTCTCTCTTACTCACAGTTGGGCCATGTATTATTAACAAGTTAATG 7004
628 laPheValArgGluArgValSerAlaValGlnIleMetValLeuArgGln 644
|||||
7005 CCTTCATTAGAACGAATAAGTGCCAGTCCAGATCATGTACTTACTAGACAA 7054
645 GlnTyrGln 647
|||||
7055 CAGTACCAA 7063

seq_name: /SID2/gc9data/geneseq/geneseq/NA1997.DAT:AA74811
seq_documentation_block:
ID AA74811 standard; cDNA; 8060 BP.
XX
AC AA74811;
XX
DT 11-FEB-1998 (first entry)
XX
DE Porcine retrovirus Tsukuba-1 cDNA.
XX
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW xenotransplantation; infectious; provirus; organ transplant; donor;
KW activated virus; Tsukuba-1; PCR; ss.
XX
OS Porcine retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 86..2002
FT /*tag= a
FT /note= "putative ENV protein"
FT CDS 3112..4686
FT /*tag= b
FT /note= "putative GAG protein"
FT CDS 4871..8060
FT /*tag= c
FT /note= "putative POL protein (partial)"
XX
PN W09721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US19680.
XX
PR 14-DEC-1995; 95US-0572645.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Fishman JA;

```





```

443 GluGluValValLeuAsePGLuTyArgTyrAsnArgProLysAr 459
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1358 GAAAGCAATCCTGTGAATATGACTACAGAAATCATGCACAAAGAG 1407
459 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuThrAlav 476
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1408 AGAACCATATCTCTGACACTTCTGTGATGCTCGGACTTGGAGTGGCAG 1457
476 alGlyValGlyThrGlyThrAlaAlaLeuLeuThrGlyProGlnGlnLeu 492
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1458 CAGGTGTAGGAACAGGAACAGCTGCCCTGTGCACGGGACCACAGACTA 1507
493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 509
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1508 GAAACAGGACTTGTAGTAACTTACATCGAATTTGAACAGAAGATCTCCAAGC 1557
509 aLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerG 526
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1558 CTTAGAAATATCTGTGAGTAACTTGGAGGAATCCCTTAACCTCTTATCTG 1607
526 luValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 542
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1608 AAGTAGTCTTACAGATAGAGAGGCTTAGATTTATTTCTAAAGAA 1657
543 GlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHi 559
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1658 GGAGGATTATGTAGCTTGTGAAGGAGGAATGCTGTTTATGTGGATCA 1707
559 sSerGlyAlaAlaLeuArgAspSerMetAsnLysLeuArgLysLysLeuGlu 576
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1708 TTCAGGGGCTATCAGAGACTCCATGAACAACTTAGAGAAAGGTTGGAGA 1757
576 rArgArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPhe 592
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1758 AGCGTCGAAGGAAAGAACTACTCAAGGTGGTTTCAGGGATGGTTC 1807
593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLe 609
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1808 AACAGGTCCTCTGTGGTGGCTACCCVACTTCTGCTTTAACAGGACCTT 1857
609 uValValLeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgP 626
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1858 AATAGTCTCTCTCTCTACTCACAGTTGGGCCATGTATTATTAACAAGT 1907
626 heValAlaPheValArgGluArgValSerAlaValGlnIleMetValLeu 642
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1908 TAATTGCTTCTATTAGAACGAATTAAGTGCAGTCCAGATCATGTACTT 1957
643 ArgGlnGlnTyrGln 647
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1958 AGACAAACAGTACCAA 1972

```

seq\_name: /SID52/qcgdata/geneseq/geneseq/NA2001.DAT:AAF77725

seq\_documentation\_block:

ID AAF77725 standard; cDNA; 8060 BP.

AC AAF77725;

XX AAF77725;

DT 23-MAY-2001 (first entry)

DE Tsukuba-1 cDNA.

XX Tsukuba-1; retrovirus; graft transplantation; xenotransplantation; ss.

XX Porcine retrovirus.

OS US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX

PF

```

XX 14-DEC-1995; 95US-0572645.
XX (GEHO ) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI; 2001-256211/26.
XX Assessing risk of endogenous retroviruses in clinical practice and in
XX xenotransplantation, comprises using probe sequences derived from swine
XX or miniature swine retroviral genome -
XX Claim 1; Fig 1; 127pp; English.
XX The present invention relates to a method for screening a cell or tissue
XX for the presence or expression of a retrovirus (RV), comprising
XX contacting a target nucleic acid from the cell or tissue with a second
XX nucleic acid from the present invention (e.g. the present sequence or a
XX fragment thereof). The method is useful for RV detection and to assess
XX graft transplantation risk. Screening of animals allows the elimination
XX of donors with active replication of known viruses. Inactive proviruses
XX can be detected and inactivated, allowing identification and elimination
XX of potential human pathogens derived from swine in a manner not possible
XX in the outbred human organ donor population and is important to the
XX development of human xenotransplantation.
XX Sequence 8060 BP; 2233 A; 1959 C; 2012 G; 1856 T; 0 other;
SQ

```

alignment\_scores:

Quality: 2402.00 Length: 655

Ratio: 4.312 Gaps: 9

Percent Similarity: 85.038 Percent Identity: 68.702

alignment\_block:

US-09-171-553B-10 x AAF77725 ..

Align seg 1/1 to: AAF77725 from: 1 to: 8060

```

1 MethHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
86 ATGCATCCACGTTAAAGCGCGCCACCTCCCGATTCCGGGTGGAAGCC 135
17 oLysArgLeuArgGileProLeuSerPheAlaSerIleAlaTrpPheLeuT 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
136 GAAAGACTGAAATCCCTTAAGCTTCGCCCTCCATCCGCTGCTTCTTA 185
34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
186 CTCTGCTCAATAACCTCTCAGACTAATGATGCGATAGGAGACAGCCTG 235
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
236 AACTCCCAATAACCCCTTATCTCTCACCTGGTTAATTAATGACTCCGC 285
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrp 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
286 AGGTATTAATATCAACACTCAAGGAGGCTCCCTTTAGGAACCTGGT 335
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
336 GGCCTGATCTATACGTTTGCCTCAGATCAGTTATTCTCTAGTCTGACCTCA 385
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 ...CCCCAGATATCTCCATGCTCAGGATTTTATGTTTCCCGAGACC 432
117 rGluLysGlu...LysTyrCysGlyGlySerGlyGlySerPheCysArgA 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
433 ACCAAATATGGAACATATCGGAAATCCAGAGATTTCTTTGTAAC 482
133 rgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149

```



```

FT mat_peptide /*tag= a
FT 585..2156 /*tag= b
FT /*note= "putative GAG protein"
FT CDS 2307..5744
FT /*tag= c
FT mat_peptide 2307..5741
FT /*tag= d
FT /*note= "putative POL protein"
FT CDS 5620..7536
FT /*tag= e
FT mat_peptide 5620..7533
FT /*tag= f
FT /*note= "putative ENV protein"
FT FT
XX WO9721836-A1.
XX
XX 19-JUN-1997.
XX
XX 13-DEC-1996; 96WO-US19680.
XX
XX 14-DEC-1995; 95US-0572645.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Fishman JA;
XX
XX WPI; 1997-332804/30.
XX
XX P-PSDB; AAW32096-W32098.
XX
XX New nucleic acid from porcine retroviruses - used for detecting
XX PT viruses in transplant or other tissue and for assessing risk of
XX PT transmitting infection to graft recipient
XX
XX Claim 22; Fig 3; 128pp; English.
XX
XX This cDNA sequence represents a porcine retrovirus from miniature swine
XX CC containing the putative coding regions for viral GAG, POL and ENV
XX CC proteins. This sequence and PCR fragments generated from it
XX CC (see AAT74812-T74882) can be used to screen organs for the presence of
XX CC porcine retroviruses prior to xenotransplantation. Transplantation can
XX CC increase the likelihood of retroviral activation if intact and
XX CC infectious proviruses are present. The porcine retroviral sequence can be
XX CC used to generate probes to determine the level (e.g. copy number) of
XX CC intact (i.e. potentially replicating) porcine provirus sequences in a
XX CC strain of xenograft transplantation donors. It can be used to detect
XX CC mutations, genetic lesions or viral recombinants and also to determine
XX CC the histological localisation of activated retrovirus. Using Polymerase
XX CC Chain Reaction DNA Quantitation (PQ) on blood mononuclear cells,
XX CC infectivity titration and susceptibility testing can be performed.
XX CC Ultimately animal donors without intact porcine retroviral sequences or a
XX CC lower copy number of viral elements could be selected.
XX
XX Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

```

```

alignment_scores:
  Quality: 2395.00      Length: 655
  Ratio: 4.292          Gaps: 10
  Percent Similarity: 85.191  Percent Identity: 68.855

```

alignment\_block:

US-09-171-553B-10 x AAT74884 ..

Align seg 1/1 to: AAT74884 from: 1 to: 7892

```

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17
|||||
5620 ATGCATCCACGTTAAACCGGGCCACCTCCCGATTCCGGGTGGAAGCC 5669
|||||
17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlarppHeLeuT 34
|||||
5670 GAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTTCCTTA 5719
|||||

```

```

34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50
|||||
5720 CTCTGTCAATAACCTCTCAGACTAATGGTATGCCATAGAGACACCGCTG 5769
|||||
51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67
|||||
5770 AACTCCCATAAACCTTATCTCTCACCTGGTTAATTAATTAATTAATTA 5819
|||||
67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84
|||||
5820 AGGTATTAAATATCAACAACACTCAAGGGGAGGCTCCTTTAGGAACCTG 5869
|||||
84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
|||||
5870 GGCCTGATCTATACGTTTGGCTTCAGATCAGTATTCTTCTAGTCTGAC 5919
|||||
101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117
|||||
5920 ...CCCCAGATATCCTCATGCTCACGATTATTTATGTTGCCCGAGGACC 5966
|||||
117 rGluLysGlu...LysTyrCysGlyGlySerGlyGluSerPheCysArgA 133
|||||
5967 ACCAAATAATGGAACCAATTCGGAAATCCCGAGAGATTCTTTGTAAAC 6016
|||||
133 rGTPSerCysValThrSerAsnAspGlyAspTrpLysTrpProIleSer 149
|||||
6017 AATGGAAGTGTAACTCTAATGATGATATTGGAATATGCCCAACCTCT 6066
|||||
150 LeuGlnAspArgValLysPheSerPheValAsn.....SerGlyProGl 164
|||||
6067 CAGCAGGATAGGTAAAGTTTCTTATGTCACACCTATACCACTCTGG 6116
|||||
164 yLysTyrLysMetMetLysLeuTyrLysAspLysSer.....CysSerP 179
|||||
6117 ACAATTTAATTAACCTGACCTGGATTAGAACTGGAAGCCCAAGTCTCTC 6166
|||||
179 roSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGln 195
|||||
6167 CTTGACAGCTAGATTACCTAAATAATAGTTTCACTGAGAAGGAAACAA 6216
|||||
196 GluAsnIleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTy 212
|||||
6217 GAAATAATCTCTAAATGGTAAATGGTATGCTTGGGAAATGGTATAT... 6264
|||||
212 rLysTyrGlyGlyGlyAla.....GlySerThrLeuThrIleArgL 226
|||||
6265 ...TATGGAGGCTCGGGTAAACACACAGGCTCCATTCTTAACATATTCGC 6310
|||||
226 euArgIleGluThrGlyThrGluProValAlaMetGlyProAspLys 242
|||||
6311 TCAAAATAAACCCAG...CTGGAGGCTCCATGGCTATAGACCAAAATACG 6357
|||||
243 ValLeuAlaGluGlnGlyProProAlaLeuGluProProHisAsnLeuPr 259
|||||
6358 GTCTTGAGGGGTCAAGAGACCCCAACCAAGGACCA..... 6393
|||||
259 oValProGlnLeuThrSerLeuArgProAspIleThrGlnProProSerA 276
|||||
6394 .....GGACCATCTCTCTA 6406
|||||
276 snSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
|||||
6407 ACATAACTTCTGGATCAGACCCCACTGAGTCT.....AACAGC..... 6444
|||||
293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaph 309
|||||
6445 ...ACGACTAAATGGGGGCAAACTTTTAGCTCATCCAGGAGGCTTT 6491
|||||
309 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuC 326
|||||
6492 TCAAGCTCTTAACTCCACGACTCCAGAGGCTACCTCTCTCTTGTGGCTAT 6541
|||||

```

326 ysLeuSerSerGlyProProTyrTyrGluGlyMetAlaLysGluArgLys 342  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6542 GCTTACGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAAAGAGGAAA 6591  
  
343 PheAsnValThrLysGluHisArgAsnGlnCysThrTyrGlySerArgAs 359  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6592 TTCATGTCGACAAAGAACATAGACCAATGCATATGGGGATCCCAAAA 6641  
  
359 nLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysA 376  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6642 TAAGCTTACCTTACTAGGTTTCTGGAAAAGGCACCTGCATAGGAAAG 6691  
  
376 laProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6692 TTCGCCCATCCCAACACCTTTTGAACCACTCAAGCCCTTAAATCAA 6741  
  
393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTyrTyrAlaC 409  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6742 ACCTCTGAGAGTCATATCTGTACTCTGGTTATGACAGGTGGTGGCATG 6791  
  
409 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerL 426  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6792 TAATACTGGATTAAACCCCTGTGTCTCCACCTTGGTTTAAACCAACTA 6841  
  
426 ysAspLeuCysValMetValGlnIleValProArgValTyrThrHisPro 442  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6842 AAGATTTTTCATATGTCGCAAAATGTTCCCGAGTGTATTAATATATCC 6891  
  
443 GluGluValValLeuAspGluTyrAspTyrArgTyrAsnArgProLysAr 459  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6892 GAAAGAGCAATCTTGATGAATATGACTACAGAAATCATCGACAAAGAG 6941  
  
459 gGluProValSerLeuThrLeuAlaValMetLeuGlyLeuGlyThrAlav 476  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6942 AGAACCCATATCTGCACACTTGTGTGTGATGCTCGGACTTGGAGTGCAG 6991  
  
476 alGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
6992 CAGGTGTAGAACAGACAGCTGCGCTGGTCACGGGACCCACAGCAGCTA 7041  
  
493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl 509  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7042 GAAACAGGACTTAGTAACCTACATCGAATGTGTAACAGAGATCTCCAAG 7091  
  
509 alLeuLysGluSerValSerAsnLeuGluSerLeuThrSerLeuSerG 526  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7092 CCTAGAAAAATCTGCAGTAACCTGGAGGAATCCCTAACCTCTTATCTG 7141  
  
526 luValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuArgGlu 542  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7142 AGTAGTCTCAGAAATAGAGAGGTTAGATTATTTATTCTAAAGAA 7191  
  
543 GlyGlyLeuCysAlaAlaLeuLysGluGlyCysPheTyrValAspH 559  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7192 GGAGGATTATGTGTAGCTTGAAGGAGGAATGCTGTTTTATGTGATCA 7241  
  
559 sSerGlyAlaIleArgAspSerMetAsnLysLeuArgLysLysLeuGluA 576  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7242 TTCAGGGGCCATCAGAGACTTCCATGAACAGCTTAGAAGAGGTGGAGA 7291  
  
576 rgArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPhe 592  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7292 AGCGTCGAAGGAAAGAAAGAACTACTCAAGGGTGGTTGAGGGATGGTTC 7341  
  
593 AsnArgSerProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLe 609  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7342 AACAGGCTCTTGTGGTGGCTACCTACTTTCTGCTTTAAACAGGACCCCT 7391  
  
609 uValValLeuLeuLeuThrValGlyProCysLeuLeuLeuAsnArgP 626  
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
7392 AATAGTCTCTCTCTGTTACTCACAGTGGGCCATGTATTATACAAGT 7441  
  
626 heValAlaPheValArgGluArgValSerAlaValIleMetValLeu 642

7442 TAATTGCTTCAATTAGAAAGCAATAAGTGCAGTCCAGATCATGGTACTT 7491

643 ArgGlnGlnTyrGln 647

7492 AGACAACAGTACCAA 7506

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF77727

seq\_documentation\_block:

ID AAF77727 standard; DNA; 8132 BP.

XX AAF77727;

XX 23-MAY-2001 (first entry)

XX Nucleotide sequence of a retrovirus found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine;  
KW ss.  
XX Unidentified.  
XX US6190861-B1.  
XX 20-FEB-2001.  
XX 13-DEC-1996; 96US-0766528.  
XX 14-DEC-1995; 95US-0572645.  
XX (GEHO ) GEN HOSPITAL CORP.  
XX Fishman JA;  
XX WPI: 2001-256211/26.  
XX P-PSDB; AAB73285, AAB73286, AAB73287.

XX Assessing risk of endogenous retroviruses in clinical practice and in  
XX xenotransplantation, comprises using probe sequences derived from swine  
XX or miniature swine retroviral genome -  
XX Claim 1; Fig 3; 127pp; English.  
XX The present invention relates to a method for screening a cell or tissue  
XX for the presence or expression of a retrovirus (RV), comprising  
XX contacting a target nucleic acid from the cell or tissue with a second  
XX nucleic acid from the present invention (e.g. the present sequence or a  
XX fragment thereof). The method is useful for RV detection and to assess  
XX graft transplantation risk. Screening of animals allows the elimination  
XX of donors with active replication of known viruses. Inactive proviruses  
XX can be detected and inactivated, allowing identification and elimination  
XX of potential human pathogens derived from swine in a manner not possible  
XX in the outbred human organ donor population and is important to the  
XX development of human xenotransplantation.

XX Sequence 8132 BP; 2248 A; 1977 C; 2037 G; 1870 T; 0 other;

alignment\_scores:

Quality: 2395.00

Ratio: 4.292

Length: 655

Gaps: 10

Percent Similarity: 85.191

Percent Identity: 68.855

alignment\_block:

US-09-171-553B-10 x AAF77727 ..

Align seg 1/1 to: AAF77727 from: 1 to: 8132

1 MethisProThrLeuSerTyrArgHisLeuProThrArgGlyGlyGluPr 17

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

5620 ATGATCCCACTTAAACCGGCCACCTCCCGATTCGGGGTGGAAAGCC 5669



7392 AATAGTCCTCCCTGTTACTACACAGTTGGCCATGTTATTATTAACAAGT 7441  
 626 heValAlaPheValArgValSerAlaValGlnIleMetValLeu 642  
 ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 7442 TAATTGCTTCATTAGAGAACGAATAAGTGCAGTCCAGATCATGTACTT 7491

643 ArgGlnGlnTyrGln 647  
 |||||||  
 7492 AGACAACAGTACCAA 7506

seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ51087

seq\_documentation\_block:

ID: AAZ51087 standard; DNA: 2000 BP.

XX AAZ51087;

XX 05-JUN-2000 (first entry)

XX Contiguous DNA of porcine endogenous retrovirus-D (PERV-D) env region.

XX Porcine endogenous retrovirus; PERV-D; virucide; prevention; vaccine;  
 KW pig; diagnosis; infection; xenotransplantation; antibody; env region; ds.

XX OS Sus scrofa.

XX WO200011187-Al.

XX PD 02-MAR-2000.

XX PF 18-AUG-1999; 99WO-US19053.

XX PR 18-AUG-1998; 98US-0097015.

XX PA (BIOT-) BIO TRANSPLANT INC.

XX PI Banerjee PT, Patience C, Andersson GK;

XX DR WPI; 2000-224704/19.

XX PT Porcine retroviral PERV-D polypeptides for diagnosing porcine  
 PT retroviral infections in humans after xenotransplantation -

XX PS Example 4; Fig 11; 119pp; English.

XX CC The present DNA is the contiguous sequence of porcine endogenous  
 CC retrovirus-D (PERV-D) env region. It is compiled from the DNA fragments  
 CC obtained from the 5' end, extended proline rich region and 3' end of  
 CC PERV-D env region. It is isolated from the porcine kidney, PK15 cell  
 CC (ATCC No.CCL-33). PERV-D DNA has 79% homology to a portion of PERV-C and  
 CC has virucidal activity. PERV-D sequence is useful for prevention or  
 CC diagnosis of infection of human tissues by porcine retroviruses after  
 CC xenotransplantation procedures. PERV-D polypeptide may be used to produce  
 CC specific antibodies, that can be administered as vaccines to create  
 CC passive immunity. The DNA can be used as a hybridisation probe or primer  
 CC for isolation purposes.

XX SQ Sequence 2000 BP; 667 A; 468 C; 366 G; 499 T; 0 other;

alignment\_scores:

Quality: 2169.00 Length: 652  
 Ratio: 4.032 Gaps: 11  
 Percent Similarity: 82.515 Percent Identity: 64.571

alignment\_block:

US-09-171-553b-10 x AAZ51087 ..

Align seg 1/1 to: AAZ51087 from: 1 to: 2000

1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluPr 17  
 ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 43 ATACATCCAGGTTAAGCGCGGCCACCTCCTGATTCAGGCGTGAAGGCC 92

17 oLysArgLeuArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuT 34  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 93 AAAAAGACTAAAAATCCCTTAAGCTTCGCTCCATCAGTGTCTCTTA 142  
 34 hrLeuThrIleThrProGlnAlaSerSerLysArgLeuIleAspSerSer 50  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 143 CTCTGTCATAAACCCTTCAGACTAATGGTATGCACATAGGAGACAGCTG 192  
 51 AsnProHisArgProLeuSerProThrTrpLeuIleIleAspProAspTh 67  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 193 AACTCCCATAAACCTTATCTCTGACCTGGTTAATTACTGACTCTGCAC 242  
 67 rGlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpT 84  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 243 AGGTATTATATCCAGCGCTCGAGGGAGGCTCCTTTAGAAAACCTGGT 292  
 84 rpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 293 GGCTGTATCTATATGCTGCTCAGATCAGTCAATCTCTAGCTGACCTCA 342  
 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyTh 117  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 343 ACC...CCAGATATCCTCGCTGCTTACGGATTTATGTTGCCACGACC 389  
 117 rGluLysGlu...LysTyrCysGlyGlySerGlyGluSerPheCysArgA 133  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 390 ACCAAATAATGGAAACACATATGGAAATCCTAGAGATTTCTTTTACA 439  
 133 rGTPSerCysValThrSerAsnAspGlyAspTrpLysTrpProLysSer 149  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 440 AATGGAGCTGTGAACCTCTAATGATGGAATCGAAAATGGCCCACTCT 489  
 150 LeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysT 166  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 490 CTGAGGATAGGGTAAGCTTTTCTTATGTCAAC...CCCAATAACCA 533  
 166 rLysMetMetLysLeu...TyrLysAspLysSer...CysSerProSerA 181  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 534 CCGGACCTGGAAACGTCATACAGGTACCATCTCGGTGTTTCCCTCAG 583  
 181 spLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsn 197  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 584 ACCTAGATTATCTAAATAAGTTTCACCGAAAAAACAAGAAAT 633  
 198 IleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLys.. 213  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 634 ATCCTAAATGGATAAATGGTATGCTCTGGGGAATAATATATTATACAG 683  
 214 TyrGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIleGluT 230  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 684 TTGGGACAGACACACCGGCTCCATTCTAACCATCCGACTTAAATA...A 730  
 230 hrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGlu 246  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 731 GCCAGCTAGAGCTCCCAATGGCTATAGGCGGAATACGGTCTTAACGGGT 780  
 247 GlnGlyProProAlaLeuGluProProHisAsnLeuProValProGlnLe 263  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 781 CAAAGAACCACCAACCGAGGACCATCTCTGATATA..... 816  
 263 uThrSerLeuArgProAspIleThrGlnProProSerAsnSerThrThrG 280  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 817 ....ACTTCTAAATTAGACCCCACTAG.....TCTAACAGACGACACT. 855  
 280 lyLeuIleProThrAsnThrProArgAsnSerProGlyValProValLys 296  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 856 .....AAA 858  
 297 ThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIleAs 313  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 859 ACGGGACAAAACCTTTTATGCTCATCCAGGGAGCTTTTCAAGCTCTAA 908







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2002, 03:05:27 ; Search time 624.99 Seconds  
(without alignments)  
27.435 Million cell updates/sec

Title: US-09-171-553B-13

Perfect score: 20

Sequence: 1 gatgctctcctgcctt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_1101.\*

1:	/SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AAV09704	Porcine retrovirus
2	20	100.0	18	AAV09698	Porcine retrovirus
3	20	100.0	22	AAV77726	Defective retrovir
4	20	100.0	18	AAV74883	Porcine retrovirus
5	20	100.0	18	AAV09699	Porcine retrovirus
6	20	100.0	18	AAV09700	Porcine retrovirus
7	18.4	92.0	3482	AAV82749	Porcine retrovirus
8	18.4	92.0	4402	AAV82749	Porcine retrovirus
9	18.4	92.0	4918	AAV82749	Porcine retrovirus
10	18.4	92.0	6076	AAV82749	Porcine retrovirus
11	18.4	92.0	6076	AAV82749	Porcine retrovirus

ALIGNMENTS

12	18.4	92.0	7362	22	AAV67019	PERV env protein c
13	18.4	92.0	7873	22	AAV67023	PERV env protein c
14	18.4	92.0	7892	18	AAV74884	Miniature swine re
15	18.4	92.0	8060	18	AAV74811	Porcine retrovirus
16	18.4	92.0	8060	22	AAV77725	Tsukuba-1 CDNA. P
17	18.4	92.0	8132	22	AAV77727	Nucleotide sequenc
18	17.4	87.0	300	21	AAV00027	Human colon cancer
19	17.4	87.0	773	21	AAV00252	Human colon cancer
20	16.8	84.0	335	22	AAV123041	Human colon cancer
21	16.8	84.0	335	22	AAV148346	Probe #12974 for g
22	16.8	84.0	335	22	AAV108699	Probe #8690 used t
23	16.8	84.0	578	22	AAV19274	Probe #9207 for ge
24	16.8	84.0	578	22	AAV14443	Probe #13129 used
25	16.8	84.0	926	21	AAV77321	Human OREF ORF2876
26	16.8	84.0	38186	22	AAV33028	Human METH1 relate
27	16.8	84.0	38186	22	AAV39085	AC004449 CDNA clon
28	16.8	84.0	72928	20	AAV18355	Human ASTH1J 5' ge
29	16.8	84.0	72928	21	AAV08253	Human ASTH1J 5' ge
30	16.4	82.0	2331	21	AAV08583	Human cytoskeleton
31	16	80.0	23	18	AAV74822	Porcine retrovirus
32	16	80.0	23	18	AAV74823	Porcine retrovirus
33	16	80.0	23	22	AAV77738	Porcine retrovirus
34	16	80.0	23	22	AAV77739	PCR primer #11. U
35	16	80.0	2035	15	AAV66537	Sequence encoding
36	16	80.0	2401	19	AAV32997	Rat jejunal concen
37	15.8	79.0	519	21	AAV33612	Arabidopsis thalia
38	15.8	79.0	1660	21	AAV49701	Arabidopsis thalia
39	15.8	79.0	1667	21	AAV43591	Arabidopsis thalia
40	15.8	79.0	10723	18	AAV49303	CDNA encoding poly
41	15.8	79.0	10723	18	AAV49303	CDNA sequence enco
42	15.8	79.0	13633	14	AAV04862	VAC-beta (genomic
43	15.4	77.0	249	22	AAV65932	C glutamicum codin
44	15.4	77.0	480	21	AAV87325	Rat hepatocyte car
45	15.4	77.0	1153	13	AAV30680	Sequence of rat GA

RESULT 1

ID	AAV09704	standard; DNA: 20 BP.
AC	AAV09704;	
XX		
XX		
DT	19-MAR-1998	(first entry)
XX		
DE	Porcine retrovirus DNA PCR primer 3.	
XX		
KW	Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;	
KW	vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic;	
KW	PCR primer; ss.	
XX		
OS	Synthetic.	
OS	Porcine retrovirus.	
XX		
PN	WO9740167-A1.	
XX		
PD	30-OCT-1997.	
XX		
PF	18-APR-1997;	97WO-GH01087.
XX		
PR	10-FEB-1997;	97GB-0002668.
PR	19-APR-1996;	96GB-0008164.
XX		
PA	(IMUT-) IMUTRAN LTD.	
XX	(ONE-) ONE BIOTECH LTD.	
XX	Galbraith DN, Haworth C, Lees GM, Smith KT;	
XX	WPI; 1997-535851/49.	
PT	Polynucleotide encoding porcine retrovirus expression product -	

PT useful to develop products for use in vaccines, diagnosis and  
PT xeno-transplantation

XX Example 7; Page 29; 69pp; English.

XX PCR primers AAV09704 and AAV09705 are used in the amplification of the  
CC porcine retrovirus (POEV). The encoded proteins within this retrovirus  
CC e.g. the virion core polypeptide (GAG), polymerase (POL) and envelope  
CC protein (ENV) can be used to develop viral vaccines, antisense nucleic  
CC acids, ribozymes and other antiviral agents. They can also be used in  
CC xeno-transplantation technology and as diagnostic tools.

XX Sequence 20 BP; 1 A; 7 C; 5 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20  
1 gatgctctctgccccttg 20

RESULT 2

AAV09698  
ID AAV09698 standard; DNA: 3320 BP.

XX AAV09698;

XX 19-MAY-1998 (first entry)

DE Porcine retrovirus pol and env DNA.

XX Porcine retrovirus; POEV; POL protein; ENV protein; vaccine;  
KW diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX Porcine retrovirus.

XX Key Location/Qualifiers

FT 23..2793

FT /\*tag= a

FT /product= POL protein

FT /note= "polymerase protein"

FT 2642..3297

FT /\*tag= b

FT /product= ENV protein

FT /note= "envelope protein"

XX W09740167-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97MO-GB01087.

XX 10-FEB-1997; 97GB-0002668.

XX 19-APR-1996; 96GB-0008164.

XX (IMUT-) IMUTRAN LTD.

XX (COONE-) Q-ONE BIOTECH LTD.

XX Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WPI; 1997-535851/49.

XX Polynucleotide encoding porcine retrovirus expression product -

XX xeno-transplantation

XX Claim 4; Fig 1; 69pp; English.

XX This DNA sequence encodes the porcine retrovirus (POEV) polymerase (POL)

XX and envelope (ENV) proteins. These proteins can be used to develop viral

CC vaccines, antisense nucleic acids, ribozymes and other antiviral agents.  
CC They can also be used in xeno-transplantation technology and as  
CC diagnostic tools.

XX Sequence 3320 BP; 922 A; 844 C; 832 G; 722 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 3320;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20  
2289 gatgctctctgccccttg 2308

RESULT 3

AA77726  
ID AA77726 standard; DNA: 7333 BP.

XX AA77726;

XX 23-MAY-2001 (first entry)

DE Defective retroviral genome isolated from PK-15 cell line.

XX Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line;

XX ss.

XX Unidentified.

XX US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

XX (GEHO ) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 2001-256211/26.

XX P-PSDB; AA773282, AA773283, AA773284.

XX Assessing risk of endogenous retroviruses in clinical practice and in

XX xenotransplantation, comprises using probe sequences derived from swine

XX or miniature swine retroviral genome -

XX Claim 1; Fig 2; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue  
CC for the presence or expression of a retrovirus (RV), comprising  
CC contacting a target nucleic acid from the cell or tissue with a second  
CC nucleic acid from the present invention (e.g. the present sequence or a  
CC fragment thereof). The method is useful for RV detection and to assess  
CC graft transplantation risk. Screening of animals allows the elimination  
CC of donors with active replication of known viruses. Inactive proviruses  
CC can be detected and inactivated, allowing identification and elimination  
CC of potential human pathogens derived from swine in a manner not possible  
CC in the outbred human organ donor population and is important to the  
CC development of human xenotransplantation.

XX Sequence 7333 BP; 1984 A; 1788 C; 1885 G; 1676 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 7333;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatgctctctgccccttg 20  
|||||

DB 4442 gatgctctcctgccttg 4461

RESULT 4  
AA74883  
ID AA74883 standard; cDNA: 7393 BP.  
XX  
XX AA74883;  
XX  
XX 09-FEB-1998 (first entry)  
XX  
XX Porcine retrovirus cDNA (defective).  
XX  
XX Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
XX xenotransplantation; infectious; provirus; organ transplant; donor;  
XX activated virus; PCR; ss.  
XX  
XX Porcine retrovirus.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 598..2172  
XX /tag= a  
XX mat\_peptide 598..2169  
XX /tag= b  
XX /note= "putative GAG protein"  
XX CDS 2320..4737  
XX /tag= c  
XX /note= "putative POL coding region (partial) as  
XX described in the specification"  
XX mat\_peptide 2320..3522  
XX /tag= d  
XX /note= "putative POL protein (partial)"  
XX mat\_peptide 3516..4328  
XX /tag= e  
XX /note= "putative POL protein (partial)"  
XX mat\_peptide 4332..4748  
XX /tag= f  
XX CDS 4738..6725  
XX /tag= g  
XX /note= "putative ENV coding region (partial) as  
XX described in the specification"  
XX mat\_peptide 4752..6722  
XX /tag= h  
XX /note= "ENV protein (partial)"  
XX  
XX WO9721836-A1.  
XX  
XX 19-JUN-1997.  
XX  
XX 13-DEC-1996; 96WO-US19680.  
XX  
XX 14-DEC-1995; 95US-0572645.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
XX  
XX Fishman JA;  
XX  
XX WPI; 1997-332804/30.  
XX P-PSDB; AAW32091-W32095.  
XX  
XX New nucleic acid from porcine retroviruses - used for detecting  
XX PT viruses in transplant or other tissue and for assessing risk of  
XX PT transmitting infection to graft recipient  
XX  
XX Claim 16; Fig 2; 128pp; English.  
XX  
XX This cDNA sequence represents a defective purified swine retrovirus  
XX CC found in PK-15 cells containing the putative coding regions for viral  
XX CC GAG, POL and ENV proteins. There are a few in frame stop codons and  
XX CC apparent frame shifts in the given coding sequence which alter features  
XX CC of the translation. This sequence and PCR fragments generated from the  
XX CC sequence (see AA74812-774882) could be used to screen organs for the  
XX CC presence of porcine retroviruses prior to xenotransplantation.

CC Transplantation can increase the likelihood of retroviral activation if  
CC CC intact and infectious proviruses are present. The porcine retroviral  
CC CC sequence can be used to generate probes to determine the level (e.g.  
CC CC copy number) of intact (i.e. potentially replicating) porcine provirus  
CC CC sequences in a strain of xenograft transplantation donors. It can be  
CC CC used to detect mutations, genetic lesions or viral recombinants and to  
CC CC determine the histological localization of activated retroviruses. Using  
CC CC polymerase chain reaction DNA quantitation (PQ) on blood mononuclear  
CC CC cells, infectivity titration and susceptibility testing can be  
CC CC performed. Ultimately animal donors without intact porcine retroviral  
CC CC sequences or a lower copy number of viral elements could be selected.  
XX  
XX Sequence 7393 BP; 2004 A; 1801 C; 1896 G; 1692 T; 0 other.

Query Match 100.0%; Score 20; DB 18; Length 7393;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatgctctcctgccttg 20  
DB 4442 gatgctctcctgccttg 4461

RESULT 5  
AAV09699  
ID AAV09699 standard; DNA: 8196 BP.  
XX  
XX AAV09699;  
XX  
XX 19-MAY-1998 (first entry)  
XX  
XX Porcine retrovirus DNA encoding, GAG, POL and ENV.  
XX  
XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein;  
XX KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.  
XX  
XX Porcine retrovirus.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 576..2126  
XX /tag= a  
XX /product= GAG protein  
XX /note= "viral core polypeptide"  
XX CDS 2143..5733  
XX /tag= b  
XX /product= POL protein  
XX /note= "polymerase peptide sequence as given in  
XX specification"  
XX CDS 5606..7576  
XX /tag= c  
XX /product= ENV protein  
XX /note= "envelope protein"  
XX  
XX WO9740167-A1.  
XX  
XX 30-OCT-1997.  
XX  
XX 18-APR-1997; 97WO-GB01087.  
XX  
XX 10-FEB-1997; 97GB-0002668.  
XX 19-APR-1996; 96GB-0008164.  
XX  
XX (IMUT-) IMUTRAN LTD.  
XX (QONE-) Q-ONE BIOTECH LTD.  
XX  
XX Galbraith DN, Haworth C, Lees GM, Smith KT;  
XX WPI; 1997-535851/49.  
XX  
XX Polynucleotide encoding porcine retrovirus expression product -  
XX PT useful to develop products for use in vaccines, diagnosis and  
XX PT xeno-transplantation

XX PS Claim 4; Fig 2; 69pp; English.

XX CC This DNA sequence encodes the porcine retrovirus (POEV) virion core

CC CC polypeptide (GAG), polymerase (POL) and envelope (ENV) proteins. These

CC CC proteins can be used to develop viral vaccines, antisense nucleic acids,

CC CC ribozymes and other antiviral agents. They can also be used in

CC CC xeno-transplantation technology and as diagnostic tools.

XX XX

SQ Sequence 8196 BP; 2165 A; 2061 C; 2147 G; 1820 T; 3 other;

Query Match 100.0%; Score 20; DB 18; Length 8196;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatgctctcctcctccttg 20

Db 5226 gatgctctcctcctccttg 5245

RESULT 6

AAV09700 AAV09700 standard; DNA; 8209 BP.

XX AC AAV09700;

XX DT 20-MAY-1998 (first entry)

XX XX Porcine retrovirus DNA encoding the LTR and GAG, POL and ENV proteins.

XX DE Porcine retrovirus; POEV; POL protein; ENV protein; GAG protein;

KM virion core polypeptide; polymerase protein; envelope protein;

KW vaccine; diagnosis; xenotransplantation; prophylactic; therapeutic; ds.

XX OS Porcine retrovirus.

XX FH Key Location/Qualifiers

FT misc\_feature 1..61

FT /tag= a

FT /note= "R LTR domain"

FT 1..588

FT /tag= b

FT 62..143

FT /tag= c

FT /note= "U5 LTR domain"

FT 588..2162

FT /tag= d

FT /product= GAG protein

FT /note= "viral core polypeptide"

FT 2163..5747

FT /tag= e

FT /product= POL protein

FT /note= "polymerase peptide"

FT 5620..7590

FT /tag= f

FT /product= ENV protein

FT /note= "envelope protein"

FT 7591..8209

FT /tag= g

FT 7638..8106

FT /tag= h

FT /note= "U3 LTR domain"

FT 7643..7686

FT /tag= i

FT /note= "Oct-1, c-Myp, LfF-1, E47 and ETS-1"

FT 7719..7739

FT /tag= j

FT /note= "Binding site for transcription factors

FT 7744..7754

FT /tag= k

FT protein\_bind

FT FT protein\_bind

FT /note= "Binding site for transcription factor AP-1/TR"

FT 7773..7785

FT /tag= l

FT /note= "Binding site for transcription factor ETS-1/GATA"

FT 7793..7834

FT /tag= m

FT /note= "Binding site for transcription factors

FT 7841..7851

FT /tag= n

FT /note= "Binding site for transcription factor AP-1"

FT 7873..7883

FT /tag= o

FT /note= "Binding site for transcription factor AP-1"

FT 7885..7914

FT /tag= p

FT /note= "Binding site for transcription factors

FT 7928..7953

FT /tag= q

FT /note= "Binding site for transcription factors E47

FT 7964..7968

FT /tag= r

FT 7985..7997

FT /tag= s

FT /note= "Binding site for transcription factor GATA"

FT 8003..8007

FT /tag= t

FT 8024..8036

FT /tag= u

FT /note= "Binding site for transcription factor GATA"

FT 8042..8046

FT /tag= v

FT 8049..8060

FT /tag= w

FT /note= "Binding site for transcription factor

FT 8075..8090

FT /tag= x

FT 1..61

FT /tag= y

FT /note= "R LTR domain"

FT 8127..8135

FT /tag= z

FT /note= "PolyA downstream element"

FT 8161..8166

FT /tag= aa

XX PN WO9740167-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-GB01087.

XX PR 10-FEB-1997; 97GB-0002668.

XX PA 19-APR-1996; 96GB-0008164.

XX PA (IMDT-) IMOTRAN LTD.

XX PI (COMET-) Q-ONE BIOTECH LTD.

XX PI Galbraith DN, Haworth C, Lees GM, Smith KT;

XX WFI, 1997-535851/49.

DR P-PDB; AAM39271; AAM39272; AAM39273.

XX

PT Polynucleotide encoding porcine retrovirus expression product -

PT useful to develop products for use in vaccines, diagnosis and

PT xeno-transplantation

XX PS Claim 4; Fig 3; 69pp; English.

XX XX This DNA sequence encodes the porcine retrovirus (POEV) virion core



XX XX Xenotransplantation; infectious agent; vaccine; ds.  
KW XX Porcine endogenous retrovirus.  
OS XX WO200071726-A1.  
XX XX 30-NOV-2000.  
XX PD 24-MAY-2000; 2000WO-US14296.  
XX PF 24-MAY-1999; 99US-0135631.  
XX PR (MAYO-) MAYO MEDICAL VENTURES.  
XX PA  
XX PI Federspiel MJ;  
XX DR WPI: 2001-032041/04.  
XX PT Inhibiting or preventing infectious agent transmission in mammalian  
XX PT transplant recipients, by introducing recombinant DNA comprising DNA  
XX PT encoding extracellular proteins of the agent into donor cells, such as  
XX PT swine cells -  
XX PS Claim 16; Page 109-111; 144pp; English.  
XX CC The present invention provides a method to prevent the transmission of  
XX CC infectious agents during xenotransplantation. This involves introducing  
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
XX CC the infectious agent, and then introducing these cells into the  
XX CC transplant recipient.  
SQ Sequence 4918 BP; 1382 A; 1173 C; 1183 G; 1180 T; 0 other;  
  
Query Match 92.0%; Score 18.4; DB 22; Length 4918;  
Best Local Similarity 95.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gatgctctctgccccttg 20  
||| ||||| ||||| ||||| |||||  
DB 1804 gatagctctctgccccttg 1823  
  
RESULT 10  
AAC67021  
ID AAC67021 standard; DNA: 6076 BP.  
XX AC AAC67021;  
XX DT 27-MAR-2001 (first entry)  
XX DE PERV env protein coding sequence SEQ ID NO: 21.  
XX KW Xenotransplantation; infectious agent; vaccine; ds.  
XX KM Xenotransplantation; infectious agent; vaccine; ds.  
XX OS Porcine endogenous retrovirus.  
XX OS WO200071726-A1.  
XX PN 30-NOV-2000.  
XX PD 24-MAY-2000; 2000WO-US14296.  
XX PF 24-MAY-1999; 99US-0135631.  
XX PR (MAYO-) MAYO MEDICAL VENTURES.  
XX PA  
XX PI Federspiel MJ;  
XX DR WPI: 2001-032041/04.  
XX PT Inhibiting or preventing infectious agent transmission in mammalian

PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -  
PS Claim 16; Page 107-109; 144pp; English.  
XX CC The present invention provides a method to prevent the transmission of  
XX CC infectious agents during xenotransplantation. This involves introducing  
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
XX CC the infectious agent, and then introducing these cells into the  
XX CC transplant recipient.  
SQ Sequence 6076 BP; 1613 A; 1512 C; 1609 G; 1342 T; 0 other;  
  
Query Match 92.0%; Score 18.4; DB 22; Length 6076;  
Best Local Similarity 95.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gatgctctctgccccttg 20  
||| ||||| ||||| ||||| |||||  
DB 5835 gatagctctctgccccttg 5854  
  
RESULT 11  
AAC67032  
ID AAC67032 standard; DNA: 6076 BP.  
XX AC AAC67032;  
XX DT 27-MAR-2001 (first entry)  
XX DE PERV env protein coding sequence SEQ ID NO: 32.  
XX KW Xenotransplantation; infectious agent; vaccine; ds.  
XX OS Porcine endogenous retrovirus.  
XX OS WO200071726-A1.  
XX PN 30-NOV-2000.  
XX PD 24-MAY-2000; 2000WO-US14296.  
XX PF 24-MAY-1999; 99US-0135631.  
XX PR (MAYO-) MAYO MEDICAL VENTURES.  
XX PA  
XX PI Federspiel MJ;  
XX DR WPI: 2001-032041/04.  
XX DE Inhibiting or preventing infectious agent transmission in mammalian  
XX DE transplant recipients, by introducing recombinant DNA comprising DNA  
XX DE encoding extracellular proteins of the agent into donor cells, such as  
XX DE swine cells -  
XX PS Claim 16; Page 117-119; 144pp; English.  
XX CC The present invention provides a method to prevent the transmission of  
XX CC infectious agents during xenotransplantation. This involves introducing  
XX CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
XX CC the infectious agent, and then introducing these cells into the  
XX CC transplant recipient.  
SQ Sequence 6076 BP; 1619 A; 1505 C; 1598 G; 1354 T; 0 other;  
  
Query Match 92.0%; Score 18.4; DB 22; Length 6076;  
Best Local Similarity 95.0%; Pred. No. 13;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gatgctctctgccccttg 20

Db 5835 gatagctctcctgccttg 5854

## RESULT 12

AAC67019  
ID AAC67019 standard; DNA; 7362 BP.  
XX  
XX AAC67019;  
AC  
XX  
XX 27-MAR-2001 (first entry)  
DE

PERV env protein coding sequence SEQ ID NO: 19.  
XX  
XX Xenotransplantation; infectious agent; vaccine; ds.  
XX  
XX Porcine endogenous retrovirus.

OS  
XX  
XX WO200071726-A1.  
XX  
XX

PD 30-NOV-2000.  
XX

PF 24-MAY-2000; 2000WO-US14296.  
XX

PR 24-MAY-1999; 99US-0135631.  
XX

PA (MAYO-) MAYO MEDICAL VENTURES.  
XX

PI Federspiel MJ;  
XX

DR WPI; 2001-032041/04.  
XX

XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -

XX  
XX Claim 16; Page 101-104; 144pp; English.

XX  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.

XX  
XX Sequence 7362 BP; 1997 A; 1821 C; 1881 G; 1663 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 7362;  
Best Local Similarity 95.0%; Pred. No. 13;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatagctctcctgccttg 20  
Db 5900 gatagctctcctgccttg 5919

## RESULT 13

AAC67023  
ID AAC67023 standard; DNA; 7873 BP.

XX  
XX AAC67023;  
AC  
XX

XX 27-MAR-2001 (first entry)  
DE

XX PERV env protein coding sequence SEQ ID NO: 23.  
XX

XX Xenotransplantation; infectious agent; vaccine; ds.  
XX

XX Porcine endogenous retrovirus.  
XX

XX  
XX WO200071726-A1.  
XX  
XX

PD 30-NOV-2000.  
XX

PF 24-MAY-2000; 2000WO-US14296.  
XX

PR 24-MAY-1999; 99US-0135631.  
XX

PA (MAYO-) MAYO MEDICAL VENTURES.  
XX

PI Federspiel MJ;  
XX

DR WPI; 2001-032041/04.  
XX

XX  
XX Inhibiting or preventing infectious agent transmission in mammalian  
PT transplant recipients, by introducing recombinant DNA comprising DNA  
PT encoding extracellular proteins of the agent into donor cells, such as  
PT swine cells -

XX  
XX Claim 16; Page 112-115; 144pp; English.

XX  
XX The present invention provides a method to prevent the transmission of  
CC infectious agents during xenotransplantation. This involves introducing  
CC to donor swine cells a recombinant DNA encoding a peptide fragment from  
CC the infectious agent, and then introducing these cells into the  
CC transplant recipient.

XX  
XX Sequence 7873 BP; 2200 A; 1913 C; 1962 G; 1798 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 7873;  
Best Local Similarity 95.0%; Pred. No. 13;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gatagctctcctgccttg 20  
Db 4629 gatagctctcctgccttg 4648

## RESULT 14

AAT74884  
ID AAT74884 standard; CDNA; 7892 BP.

XX  
XX AAT74884;  
AC  
XX

XX 09-FEB-1998 (first entry)  
DE

XX Miniature swine retrovirus CDNA.

XX  
XX Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
KW xenotransplantation; infectious; provirus; organ transplant; donor;

XX  
XX activated virus; PCR; ss.

XX  
XX Porcine retrovirus.

XX  
XX Key Location/Qualifiers  
FH CDS 585..2159

FT mat\_peptide 585..2156

FT /tag= a

FT /tag= b

FT /note= "putative GAG protein"

FT CDS 2307..5744

FT mat\_peptide 2307..5741

FT /tag= c

FT /tag= d

FT CDS 5620..7536

FT mat\_peptide 5620..7533

FT /tag= e

FT /tag= f

XX /note= "putative ENV protein"

XX WO9721836-A1.

XX 19-JUN-1997.

XX 13-DEC-1996; 96WO-US19680.  
 XX  
 XX 14-DEC-1995; 95US-0572645.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 XX  
 P1 Fishman JA;  
 XX  
 DR WP1: 1997-332804/30.  
 DR P-PSDB; AAW32096-W32098.  
 XX  
 XX  
 PT New nucleic acid from porcine retro:viruses - used for detecting  
 PT viruses in transplant or other tissue and for assessing risk of  
 PT transmitting infection to graft recipient  
 XX  
 PS Claim 22; Fig 3; 128pp; English.  
 XX

CC This cDNA sequence represents a porcine retrovirus from miniature swine  
CC containing the putative coding regions for viral GAG, POL, and ENV  
CC proteins. This sequence and PCR fragments generated from it  
CC (see AAT74812-774882) can be used to screen organs for the presence of  
CC porcine retroviruses prior to xenotransplantation. Transplantation can  
CC increase the likelihood of retroviral activation if intact and  
CC infectious proviruses are present. The porcine retroviral sequence can be  
CC used to generate probes to determine the level (e.g. copy number) of  
CC intact (i.e. potentially replicating) porcine provirus sequences in a  
CC strain of xenograft transplantation donors. It can be used to detect  
CC mutations, genetic lesions or viral recombinants and also to determine  
CC the histological localisation of activated retrovirus. Using Polymerase  
CC Chain Reaction DNA quantitation (PDU) on blood mononuclear cells,  
CC infectivity titration and susceptibility testing can be performed.  
CC ultimately animal donors without intact porcine retroviral sequences or a  
CC lower copy number of viral elements could be selected.

SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T; 0 other;

Query Match	92.0%;	Score 18.4;	DB 18;	Length 7892;
Best Local Similarity	95.0%;	Pred. No. 13;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	gatgctctcctg	cccttg	20
Db	5237	gatagctctcctg	cccttg	5256

RESULT 15  
2AM74911

ID AAT74811 standard; cDNA; 8060 BP.

AC AAT74811;

DT 11-FEB-1998 (first entry)

DE Porcine retrovirus Tsukuba-1 cDNA.

KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;  
KW xenotransplantation; infectious; provirus; organ transplant; donor;

05 Porcine retrovirus.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

ET

	FT	FT	FT	CDS
				/note= "putative ENV protein"
				3112..4686

FT	CDS
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

F7

```
/note= "putative POL protein (partial)"
```

XX	WO9721836-A1.
PN	
XX	
PD	19-JUN-1997.
XX	

PF	13-DEC-1996;	96WO-US19680.
XX		
PR	14-DEC-1995;	95US-0572645.

PA (GEHO ) GEN HOSPITAL CORP.

PI Fishman JA; xv

DR WPI; 1997-332804/30.

PT New nucleic acid from porcine retro:viruses - used for detecting viruses in transplant or other tissue and for assessing risk of PT

transmitting infection to graft recipient

PS Claim 1; Fig 1; 128pp; English.  
 YX

CC This sequence represents the purified porcine retroviral cDNA  
CC sequence of Tsukuba-1 and contains the putative coding region

CC proteins GAG, POL and ENV. This sequence and PCR fragments generated  
CC from the sequence (see AAT74812-T74882) could be used to screen organs

CC for porcine retroviruses prior to xenotransplantation. Transplantation  
CC can increase the likelihood of retroviral activation if intact and

CC infectious proviruses are present. The porcine retroviral sequence can be used to generate probes to determine the level (e.g. copy number) of CC

CC intact (i.e. potentially replicating) porcine provirus sequences in a  
CC strain of xenograft transplantation donors. It can be used to detect

CC mutations, genetic lesions or viral recombinants and to determine the histological localisation of activated retrovirus. Using Polymerase Chain

CC titration and susceptibility testing can be performed. Ultimately animal

CC donors without intact porcine retroviral sequences or with a lower copy  
CC number of viral elements could be selected.

Sequence 8060 BP: 2233 A: 1959 C: 2012 G: 1856 T: 0 other: 0

Query Match	92.0%;	Score 18.4;	DB 18;	Length 8060;
Best Local Similarity	95.0%;	Pred. No. 13;		

Matches	19;	conservative	0;	mismatches	1;	indels	0;	gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

1 20  
2 19  
3 18  
4 17  
5 16  
6 15  
7 14  
8 13  
9 12  
10 11  
11 10  
12 9  
13 8  
14 7  
15 6  
16 5  
17 4  
18 3  
19 2  
20 1

[illegible]

Search completed: February 24, 2002, 03:05:29  
Job time: 19495 sec





